Energy and force stepping integrators in Lagrangian mechanics

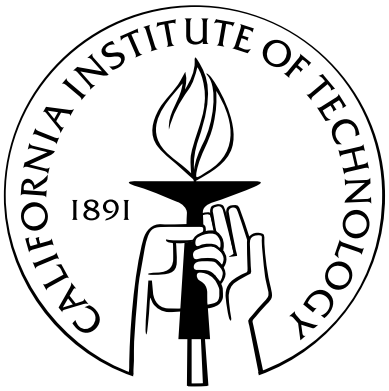
Thesis by

Marcial Gonzalez

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2011

(Defended August 20, 2010)

Acknowledgements

I would like to express my deep gratitude to Professor Michael Ortiz for giving me the opportunity to work with him and for his significant contribution to my own growth as a researcher. His wealth of ideas and broad range of scientific interests have been inspirational; his constant support and encouragement have been a driving force behind this work. I am grateful for this experience.

I would like to acknowledge Professors Chiara Daraio, Nadia Lapusta and Guruswami Ravichandran for kindly agreeing to serve on my thesis committee. Furthermore, I would like to especially acknowledge Dr. Bernd Schmidt for the insightful and enjoyable collaboration we had over the last four years, this dissertation has benefited greatly from his participation. I would also like to thank Professor Daraio for the many valuable discussions on granular crystal dynamics and Dr. Jinkyu Yang for providing experimental data used in this work.

I will always be grateful to Professors Eduardo Dvorkin and Marcela Goldschmit from the University of Buenos Aires for introducing me into the field of computational mechanics and for encouraging me to pursue a doctoral degree in the United States.

I am greatly indebted to my cherished friends whose names I will not attempt to list, but to whom I will express my heartfelt gratitude in person—as my most sincere form of appreciation. I also owe a special thanks to my past and present lab mates and fellow students for making my life at Caltech more enjoyable.

It has been a pleasure to know Lydia and Marta not only in their role of administrative and technical pillars of the group but also on a more personal level. Thank you!

Finally but most importantly, this work would not have been possible without the love and support of my parents, my sister, and of course **Gabriela**.

Abstract

The overarching goal of this thesis is to develop new numerical time integration schemes for Lagrangian mechanics that better cope with the challenges of understanding the dynamic behavior of materials. We specifically address the formulation of convergent time integration schemes that exhibit good long-term behavior—such as conferred by symplecticity and exact conservation properties—and that have the ability to automatically and asynchronously modulate the time step in different regions of the domain. We achieve these properties in a progression of three developments: (i) energy-stepping, (ii) force-stepping, and (iii) asynchronous energy-stepping integrators. These developments are based on a new method of approximation for Lagrangian mechanics, proposed in this thesis, that consists of replacing the Lagrangian of the system by a sequence of approximate Lagrangians that can be *solved exactly*. Then, energy-stepping integrators result from replacing the potential energy by a piecewise constant approximation, force-stepping integrators result from replacing the potential energy by a piecewise affine approximation, and asynchronous energy-stepping integrators result from replacing localized potential energies by piecewise constant approximations. Throughout the dissertation, the properties of these time integrators are theoretically predicted and born out by a number of selected examples of application. Furthermore, we address the challenges of understanding the propagation of solitary waves in granular crystals at low impact velocity conditions by investigating the role of energy-trapping effects with the numerical time integration schemes developed in this work.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

For the past several years, the challenges of understanding the dynamic behavior of materials have inspired the development of new experimental techniques, theoretical models and numerical methods. The overarching goal of this thesis is to develop new numerical time integration schemes for Lagrangian mechanics that cope with some of these challenges. Our sources of inspiration are those applications in material science and structural mechanics characterized by dynamical systems that exhibit, in time and space, a wide range of time scales and whose *interesting* dynamics take place in a small region of the domain that may drift in space as time progresses (e. g., shock waves, solitary waves, dynamic fracture, propagation of structural and material instabilities, many-body impact dynamics, etc.). Thus, in this dissertation, we specifically address the formulation of convergent time integration schemes that exhibit good long-term behavior—such as conferred by symplecticity and exact conservation properties—and that have the ability to automatically and asynchronously modulate the time step in different regions of the domain. It is worth noting that the formulation of convergent time integration schemes with exact conservation properties has been a longstanding goal of numerical analysis and computational mechanics. Finally, we address the challenges of understanding the propagation of solitary waves in granular crystals at low impact velocity conditions by investigating the role of energy-trapping effects with the numerical time integration schemes developed in this work.

This thesis is organized as follows. Each chapter is conceived as a separate journal article, with its own introduction, motivation, conclusions and avenues of future research. However, many of

these directions of research are carried out in full in subsequent chapters. Thus, we intend this dissertation as a progression of ideas and intertwined but distinguishable developments. Specifically, the work incorporated in and arising from this thesis is:

- *Energy-stepping integrators in Lagrangian mechanics*, Gonzalez, Schmidt, and Ortiz [24].

  We present a class of integration schemes for Lagrangian mechanics, referred to as energy-stepping integrators, that are momentum and energy conserving, symplectic and convergent. This result has been a longstanding goal of numerical analysis and computational mechanics.

  ▷ We include this work in Chapter 2.

- *Force-stepping integrators in Lagrangian mechanics*, Gonzalez, Schmidt, and Ortiz [25].

  We formulate an integration scheme for Lagrangian mechanics, referred to as the force-stepping scheme, that is symplectic, energy conserving, time-reversible and convergent with automatic selection of the time step size. The scheme also conserves approximately all the momentum maps associated with the symmetries of the system. Exact conservation of momentum maps may additionally be achieved by recourse to Lagrangian reduction.

  ▷ We include this work in Chapter 3.

- *Asynchronous energy-stepping integrators in Lagrangian mechanics*, Gonzalez, *et al.* [26].

  We develop a symplectic-energy-momentum time integrator for Lagrangian mechanics, referred to as *asynchronous energy-stepping* integrators, that has the ability to automatically and asynchronously modulate the time step in different regions of the domain, if any. The time integrator additionally accounts for the exact momentum fluxes across subdomain boundaries.

  ▷ We include this work in Chapter 4.

- *Mesoscopic approach to granular crystal dynamics*, Gonzalez, *et al.* [27].

  We present a mesoscopic approach to granular crystal dynamics, which comprises a three-dimensional finite-element model and a one-dimensional regularized contact model. The approach aims at investigating the role of vibrational-energy trapping effects in the overall dynamic behavior of one-dimensional chains under small to moderate impact velocities, and has

as its only inputs the geometry and the elastic material properties of the individual particles that form the system.

▷ We include this work in Chapter 5.

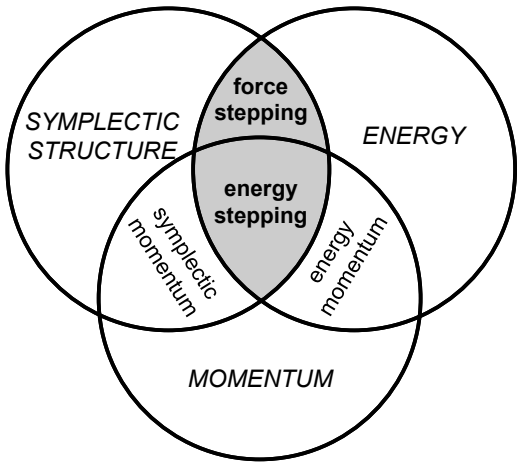


Figure 1.1: The numerical time integration schemes presented in this thesis, namely energy-stepping, force-stepping and asynchronous energy-stepping integrators, bridge the gap between two areas of active research in numerical time integration of Lagrangian mechanics: symplectic-momentum and energy-momentum integrators.

Chapter 2

Energy-stepping integrators

We present a class of integration schemes for Lagrangian mechanics, referred to as *energy-stepping integrators*, that are momentum and energy conserving, symplectic and convergent. In order to achieve these properties we replace the original potential energy by a piecewise constant, or *terraced* approximation at steps of uniform height. By taking steps of diminishing height, an approximating sequence of energies is generated. The trajectories of the resulting approximating Lagrangians can be characterized explicitly and consist of intervals of piecewise rectilinear motion. We show that the energy-stepping trajectories are symplectic, exactly conserve all the momentum maps of the original system and, subject to a transversality condition, converge to trajectories of the original system when the energy step is decreased to zero. These properties, the excellent long-term behavior of energy-stepping and its automatic time-step selection property are born out by selected examples of application, including the dynamics of a frozen Argon cluster, the spinning of an elastic cube and the collision of two elastic spheres.

2.1 Introduction

The formulation of convergent time-integration schemes with exact conservation properties has been a longstanding goal of numerical analysis and computational mechanics. In a seminal contribution, Ge and Marsden [19] showed that, in general, integrators with a fixed time step cannot simultaneously preserve energy, the symplectic structure and other conserved quantities, such as linear and angular momenta. This observation led to a dichotomy in the literature between symplectic-momentum and energy-momentum integrators. In principle, symplectic-energy-momentum integrators can be devised by adopting a space-time viewpoint and allowing for time-step adaption in order to satisfy the constraint of conservation of energy [40]. However, due to lack of solvability of the time-step optimization problem, especially near *turning points* where velocities are small [40, 51], non energy-preserving steps are inevitable in some schemes.

The design of symplectic-momentum methods can be accomplished in a systematic and natural manner by recourse to a discrete version of Hamilton's variational principle (cf., e. g., [52, 53, 58]; see also [30], Chapter VIII, and references therein) based on an approximating discrete Lagrangian. The resulting variational time integrators are, for a non-dissipative and non-forced case, symplectic and momentum preserving. However, it should be carefully noted that the momenta conserved by variational integrators are those corresponding to the discrete Lagrangian, which in general approximate—but differ from—those of the original Lagrangian. Many standard integrators, whether explicit or implicit, are indeed variational in this sense, including the widely used Newmark method [42], midpoint rule, symplectic partitioned Runge-Kutta schemes, and others (cf., e. g., Marsden and West [58] for details). Experience has shown that, while not being exactly energy conserving, constant time-step symplectic-momentum methods tend to exhibit good energy behavior over long times (see, for example, [42] and references therein), though the good energy behavior tends to degrade when variable time-steps are used [8, 20].

In general, good geometrical and invariance properties do not supply a guarantee of good accuracy and convergence of integration schemes, which must be carefully analyzed separately. In the context of linear structural systems, a powerful tool for investigating convergence is furnished by *phase-error*

analysis [3, 35, 83]. Phase error analysis aims to establish the convergence of the amplitude and frequency of oscillatory numerical solutions to the amplitude and frequency of the exact solution. Phase error analysis is a particularly powerful tool inasmuch as it establishes the convergence of solutions in a global, instead of merely local, sense. This attribute is in analogy to backward error analysis [29, 30, 71], which is also global in nature, and in contrast to other conventional methods of analysis, such as Gronwall's inequality [58], which provide local exponential bounds on discretization errors. The engineering literature on the subject of phase-error analysis relies on a case-by-case analysis of linear time-stepping algorithms. An extension of phase-error analysis to nonlinear systems may be accomplished by recourse to notions of Γ-convergence [62].

The work presented in this chapter is concerned with the formulation of time-integration schemes for Lagrangian mechanics that are symplectic and exactly conserve all the invariants of the system by construction. We achieve these desirable properties by the simple device of replacing the potential energy of the system by an approximating sequence of energies that have the same invariance properties as the original potential energy and whose trajectories can be characterized exactly. This strategy may be regarded as the reverse of that underlying backward-error analysis. Thus, whereas backward-error analysis seeks to identify a nearby Lagrangian that is solved exactly by the solutions generated by a numerical integrator, the approach followed here is to directly replace the Lagrangian by a nearby one that can be solved exactly. If the approximating Lagrangians are constructed so as to have the same symmetries as the original Lagrangian, then it follows from standard theory that all the invariants, or momentum maps, of the original system are conserved exactly by the numerical solutions. Additionally, since the numerical trajectories are exact solutions of a Lagrangian it follows that the numerical trajectories are symplectic. By generating a sequence of approximating Lagrangians that converges to the original Lagrangian in an appropriate sense, a sequence of approximating trajectories is generated that may be expected to converge to trajectories of the original system. By this simple procedure, momentum and energy conserving, symplectic and convergent time-integration schemes can be constructed.

We specifically investigate piecewise-constant approximations of the potential energy. Thus, we

replace the original potential energy by a stepwise, or *terraced*, approximation at steps of uniform height. By taking energy steps of diminishing height, an approximating sequence of energies is generated. The solutions of the approximating Lagrangians thus defined are piecewise rectilinear and can be constructed explicitly. In reference to the manner in which the potential energy is approximated and the numerical solution proceeds, we refer to the resulting integration scheme as *energy-stepping*. The rectilinear segments that comprise the energy-stepping trajectory span constant-energy surfaces and are thus of variable duration, which results in automatic time-step selection. By construction, the approximating Lagrangians are invariant under all the action groups that the original Lagrangian is itself invariant under, i.e., the approximating Lagrangians inherit all the symmetries of the original Lagrangian, and thus the energy-stepping solution conserves all the momentum maps of the original Lagrangian exactly. In addition, since the energy-stepping trajectories are exact solutions of a Lagrangian they are automatically symplectic. Furthermore, we show that, subject to a technical transversality condition, the energy-stepping trajectories converge to trajectories of the original system as the energy step decreases to zero.

The chapter is organized as follows. The energy-stepping time-integration scheme is defined in Section 2.2. The conservation properties of energy-stepping are presented in Section 2.3. In Section 2.4 we proof convergence of trajectories satisfying a transversality condition. Finally, in Section 2.5 we present selected examples of application that illustrate the convergence, accuracy and conservation properties of energy-stepping, including the dynamics of a frozen Argon cluster, the spinning of an elastic cube, and the collision of two elastic spheres. Finally, a summary and concluding remarks are collected in Section 2.6.

2.2 Energy-stepping integrators

For definiteness, we specifically consider dynamical systems characterized by Lagrangians $L : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ of the form

$$L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V(q) \tag{2.1}$$

where M is the mass matrix and V is the potential energy. Lagrangians of this form arise in a number of areas of application including structural dynamics and molecular dynamics. The trajectories of a Lagrangian system can be approximated by replacing $L(q, \dot{q})$ by an approximating Lagrangian $L_h(q, \dot{q})$ that can be solved exactly. A particular type of approximating Lagrangian is

$$L_h(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V_h(q) \tag{2.2}$$

obtained by introducing an approximation of the potential energy. In this work, we specifically investigate piecewise constant approximations of the potential energy. Thus, we replace the original potential energy V by a stepwise or *terraced* approximation V_h at steps of uniform height h, namely,

$$V_h = \text{largest piecewise-constant function with values in } h\mathbb{Z} \text{ majorized by } V \tag{2.3}$$

or, equivalently,

$$V_h(q) = h\lfloor h^{-1}V(q)\rfloor \tag{2.4}$$

where $\lfloor\cdot\rfloor$ is the floor function, i.e., $\lfloor x\rfloor = \max\{n \in \mathbb{Z} : n \leq x\}$. By taking steps of diminishing height, an approximating sequence of potential energies and Lagrangians is generated in this manner. The chief characteristics of the new systems thus obtained are that they can be solved exactly, as demonstrated in Section 2.2.1, and that they have the same symmetries as the original system, as shown in Section 2.3. Evidently, other types of approximations, such as piecewise linear interpolations of the potential energy, also result in exactly integrable approximating systems. However, a naive piecewise linear approximation breaks the symmetries of the system in general. Piecewise constant and piecewise linear approximations of the Kepler potential are shown in Figure 2.1 by way of illustration.

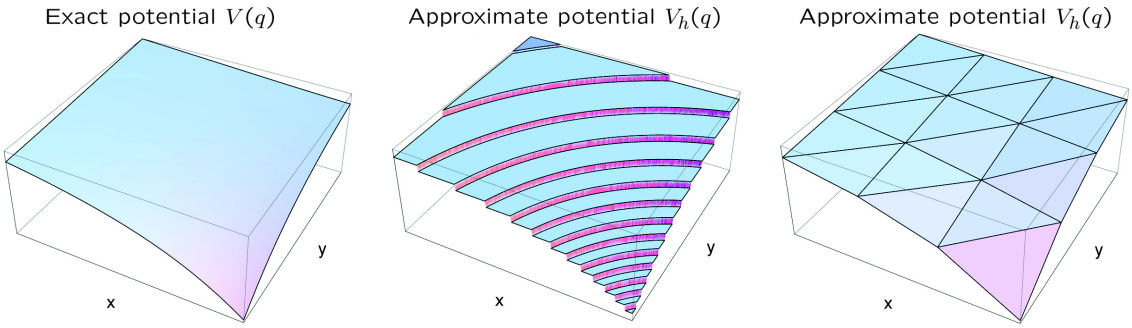


Figure 2.1: Kepler problem. Exact, piecewise constant and piecewise linear continuous approximate potential energies.

2.2.1 Computation of the exact trajectories of the approximating Lagrangian

We now proceed to compute the exact trajectories of the approximate Lagrangian $L_h(q, \dot{q})$ resulting from the piecewise constant approximation (2.3) of the potential energy. Suppose that the system is in configuration q_0 at time t_0 and in configuration q_2 at time t_2 and that during the time interval $[t_0, t_2]$ the system intersects one single jump surface Γ_1 separating two regions of constant energies V_0 and V_2, Figure 2.2. By the construction of V_h, Γ_1 is the level surface $V = V_2$ for an uphill step $V_2 = V_0 + h$, or the level surface $V = V_0$ for a downhill step, $V_2 = V_0 - h$. For simplicity, we shall further assume that V is differentiable and that all energy-level crossings are transversal, i.e.,

$$n(q_1) \cdot \dot{q}_1^- \neq 0 \tag{2.5}$$

where $\dot{q}_1^- = \dot{q}\left(t_1^-\right)$ and $n(q_1)$ is a vector normal to Γ_1 pointing in the direction of advance.

It is possible for discrete trajectories be *trapped* and become *embedded* within a jump surface for a finite time interval, e. g., when the intersection of the trajectory with the jump surface is non-transversal and the surface turns toward the trajectory in the direction of advance. In this case, the embedded trajectory segment is a geodesic of the jump surface in metric defined by the mass matrix M. However, the likelihood of non-transversal intersections resulting in the trapping of trajectories within jump surfaces is exceedingly low in practical applications and will not be given further consideration.

Under the preceding assumptions, the action integral over the time interval $[t_0, t_2]$ follows as

$$I_h = \int_{t_0}^{t_2} L_h\left(q, \dot{q}\right)\ dt = \int_{t_0}^{t_1} L_h(q, \dot{q})\ dt + \int_{t_1}^{t_2} L_h(q, \dot{q})\ dt \tag{2.6}$$

where t_1 is the time at which the trajectory intersects Γ_1. In regions where $V_h(q)$ is constant the trajectory $q(t)$ is linear in time. Therefore, the action of the system can be computed exactly and reduces to

$$I_h = (t_1 - t_0)\left\{\frac{1}{2}\left(\frac{q_1 - q_0}{t_1 - t_0}\right)^T M\left(\frac{q_1 - q_0}{t_1 - t_0}\right) - V_0\right\} + (t_2 - t_1)\left\{\frac{1}{2}\left(\frac{q_2 - q_1}{t_2 - t_1}\right)^T M\left(\frac{q_2 - q_1}{t_2 - t_1}\right) - V_2\right\} \tag{2.7}$$

where $q_1 = q\left(t_1\right)$ is constrained to be on the jump surface Γ_1. Assuming differentiability of Γ_1, stationarity of the action with respect to (t_1, q_1) additionally gives the energy conservation equation

$$\left(\frac{q_1 - q_0}{t_1 - t_0}\right)^T M\left(\frac{q_1 - q_0}{t_1 - t_0}\right) + 2V_0 = \left(\frac{q_2 - q_1}{t_2 - t_1}\right)^T M\left(\frac{q_2 - q_1}{t_2 - t_1}\right) + 2V_2 \tag{2.8}$$

and the linear momentum balance equation

$$M\frac{q_1 - q_0}{t_1 - t_0} - M\frac{q_2 - q_1}{t_2 - t_1} + \lambda n(q_1) = 0 \tag{2.9}$$

where λ is a Lagrange multiplier.

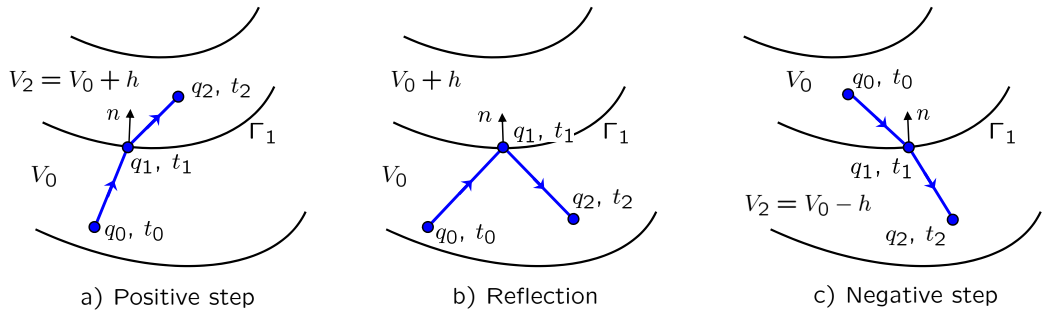


Figure 2.2: Trajectory of a system whose potential energy is approximated as piecewise constant.

In order to make a more direct connection with time-integration schemes we reformulate the problem slightly by assuming that t_0, q_0—the latter on a jump surface Γ_0 except, possibly, at the initial time—and the initial velocity

$$\dot{q}_0^+ = \dot{q}\left(t_0^+\right) = \frac{q_1 - q_0}{t_1 - t_0} \tag{2.10}$$

are known. Let t_1 and q_1 be the time and point at which the trajectory intersects the next jump surface Γ_1. We then seek to determine

$$\dot{q}_1^+ = \dot{q}\left(t_1^+\right) \tag{2.11}$$

A reformulation of eqs. (2.8) and (2.9) in terms of $\dot{q}_1^+$ gives

$$\left(\dot{q}_1^+\right)^T M \dot{q}_1^+ = \left(\dot{q}_1^-\right)^T M \dot{q}_1^- - 2\Delta V \tag{2.12}$$

$$\dot{q}_1^+ = \dot{q}_1^- + \lambda M^{-1} n(q_1) \tag{2.13}$$

where $\dot{q}_1^- = \dot{q}_0^+$ and the energy jump $\Delta V = V_h\left(q(t_1^+)\right) - V_h\left(q(t_1^-)\right)$. Next we proceed to parse through the various cases that arise in the solution of (2.12) and (2.13).

2.2.1.1 Diffraction by downhill energy step

Suppose that $\Delta V = -h$, i.e., the system decreases its energy as the trajectory crosses Γ_1. Then (2.12) becomes

$$\left(\dot{q}_1^+\right)^T M \dot{q}_1^+ = \left(\dot{q}_1^-\right)^T M \dot{q}_1^- + 2h \tag{2.14}$$

Then, the system (2.13 - 2.14) has the real solution

$$\dot{q}_1^+ = \dot{q}_1^- - \frac{\dot{q}_1^- \cdot n_1 + \sqrt{\left(\dot{q}_1^- \cdot n_1\right)^2 + 2h\, n_1^T M^{-1} n_1}}{n_1^T M^{-1} n_1} M^{-1} n_1 \tag{2.15}$$

This solution represents the diffraction, or change of direction, of the trajectory by a downhill energy step.

2.2.1.2 Diffraction by uphill energy step

Suppose now that $\Delta V = h$, i.e., the system increases its energy as the trajectory crosses Γ_1. Then (2.12) becomes

$$\left(\dot{q}_1^+\right)^T M \dot{q}_1^+ = \left(\dot{q}_1^-\right)^T M \dot{q}_1^- - 2h \tag{2.16}$$

Suppose, in addition, that

$$\left(\dot{q}_1^- \cdot n_1\right)^2 > 2h\, n_1^T M^{-1} n_1 \tag{2.17}$$

Then, the system (2.13 - 2.16) again has a real solution, namely,

$$\dot{q}_1^+ = \dot{q}_1^- - \frac{\dot{q}_1^- \cdot n_1 + \sqrt{\left(\dot{q}_1^- \cdot n_1\right)^2 - 2h\, n_1^T M^{-1} n_1}}{n_1^T M^{-1} n_1} M^{-1} n_1 \tag{2.18}$$

This solution again represents the diffraction of the trajectory by an uphill energy step when the system has sufficient initial energy to overcome the energy barrier.

2.2.1.3 Reflection by uphill energy step

Suppose now that $\Delta V = h$, i.e., the system increases its energy as the trajectory crosses Γ_1, but, contrary to the preceding case,

$$\left(\dot{q}_1^- \cdot n_1\right)^2 < 2h\, n_1^T M^{-1} n_1 \tag{2.19}$$

Then, the system (2.13 - 2.16) has no real solutions, showing that the system does not have sufficient initial energy to overcome the energy barrier. Instead, the trajectory remains within the same energy level and equation (2.12) becomes

$$\left(\dot{q}_1^+\right)^T M \dot{q}_1^+ = \left(\dot{q}_1^-\right)^T M \dot{q}_1^- \tag{2.20}$$

Then, the system (2.13 - 2.20) is solved by

$$\dot{q}_1^+ = \dot{q}_1^- - 2\frac{\dot{q}_1^- \cdot n_1}{n_1^T M^{-1} n_1} M^{-1} n_1 \tag{2.21}$$

This solution represents the reflection of the trajectory by an uphill energy step when the system does not have sufficient initial energy to overcome the energy barrier.

2.2.2 Derivation of the Hamilton principle for the discontinuous energy stepping potential by regularization

The energy stepping potential (2.4) is not continuous and the corresponding action functional (2.6) is not Gateaux-differentiable on the space of curves in phase space. In particular, the first variation of I_h is not well-defined on this space and so the notion of critical points does not have a well-defined meaning in the classical sense. However, the calculations in the previous subsection can be given a rigorous interpretation by recourse to a smooth approximation procedure. In the spirit of identifying solutions to non-smooth differential equations by suitable selection principles, we will show that the energy-stepping dynamics for fixed h is given by those solutions to (2.8) and (2.9) that are limiting trajectories of a nearby smooth dynamical system. It should be carefully noted that there are solutions to these equations other than those described in Section 2.2.1, as exemplified by a particle bouncing at an energy barrier that can be overcome energetically.

This approximation result has three important consequences:

(i) The notion of a critical point of I_h is given a rigorous definition.

(ii) The physical solutions to (2.8) and (2.9) are selected.

(iii) The existence of a nearby smooth system will prove extremely useful when investigating analytical aspects of the energy-stepping trajectories such as conservation properties (cf. Section 2.3).

Here we call a solution q to the equations (2.8) and (2.9) *physical* if and only if q behaves as described in Paragraph 2.2.1.1, 2.2.1.2 or 2.2.1.3. For later use we note that this is the case if and only if q passes an energy jump surface precisely if this is possible without violating energy conservation. As a consequence, an energy-stepping trajectory is uniquely given by $q(0)$ and $\dot{q}(0)$.

For fixed $h > 0$ consider the smooth approximation χ_ε to the stepping function $t \mapsto h\lfloor h^{-1}t \rfloor$ as given by the convolution

$$\chi(t) = \int_{\mathbb{R}} \varepsilon^{-1}\eta(\varepsilon^{-1}(t-s))\, h\lfloor h^{-1}s \rfloor \, ds, \qquad \varepsilon \ll h \tag{2.22}$$

where η denotes a standard mollifier, i.e., $\eta \in C_c^\infty(-1,1)$, $\eta \geq 0$ and $\int_{-1}^{1} \eta = 1$. Now define

$$L_{h,\varepsilon}(q,\dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V_{h,\varepsilon}(q) \quad \text{for} \quad V_{h,\varepsilon}(q) := \chi_\varepsilon(V(q)) \tag{2.23}$$

Since $L_{h,\varepsilon}$ is a smooth function, there is a well-defined classical dynamics corresponding to this Lagrangian. We will prove that, as $\varepsilon \to 0$, the critical points of the action functional

$$I_{h,\varepsilon}(q) = \int_{t_0}^{t_2} L_{h,\varepsilon}(q,\dot{q})\, dt \tag{2.24}$$

converge to the trajectories identified in Section 2.2.

More precisely, we have the following

Theorem 2.2.1 *Suppose q is an energy-stepping trajectory satisfying $\nabla V(q) \neq 0$ and $(n \cdot \dot{q}^-)^2 \notin \{0, 2hn^T M^{-1} n\}$ for $n = n(q) = \frac{\nabla V(q)}{|\nabla V(q)|}$ on every energy jump surface. Let q_ε be the smooth trajectory corresponding to $L_{h,\varepsilon}$ with the same initial conditions $(q_0, \dot{q}_0)$. Then q_ε converges to q strongly in $W^{1,p}(0,T)$ for each $T > 0$ and all $p < \infty$ and weak* in $W^{1,\infty}(0,T)$. Moreover, $\dot{q}_\varepsilon \to \dot{q}$ uniformly on $[0,T] \setminus S_r(q)$ for every $r > 0$, where $S_r(q)$ is the r-neighborhood of the jump set of $\dot{q}$.*

It is worth noting that a proof for this result can be given that does not make use of the explicit formulae for q obtained Section 2.2. Instead, we will see that limiting trajectories satisfy the discrete variational principle leading to (2.8) and (2.9) and thus give an a priori justification of the procedure employed in Section 2.2 to compute q.

Proof. First note that as $\varepsilon \to 0$, $\chi_\varepsilon(V(q)) \to h\lfloor h^{-1}V(q) \rfloor = V_h(q)$. In particular, $q_\varepsilon(t)$ is linear in time except in a small neighborhood of the energy jump surfaces. It suffices to consider the case when q meets a jump surface Γ precisely once, at time $t_1 \in (t_0, t_2)$, for example, and so we may

assume that q is a physical solution to (2.8) and (2.9).

We begin by showing that $q_\varepsilon \to \tilde{q}$ as $\varepsilon \to 0$ strongly in $W^{1,p}(t_0,t_2)$ for all $p < \infty$ (and weak* in $W^{1,\infty}(t_0,t_2)$), where $\tilde{q}$ is linear on (t_0,t_1) and (t_1,t_2). By Γ_ε we denote the neighborhood of Γ defined by

$$\Gamma_\varepsilon = \{x \in \mathbb{R}^N : V_2 - \varepsilon \leq V(x) \leq V_2 + \varepsilon\} \tag{2.25}$$

Note that as $\nabla V(q(t_1)) \neq 0$ by assumption, in the vicinity of $q(t_1)$ Γ_ε lies in an $O(\varepsilon)$-neighborhood of Γ. q_ε being a critical point of $I_{h,\varepsilon}$ satisfies the Euler-Lagrange equation

$$\ddot{q}_\varepsilon = -M^{-1}\nabla V_{h,\varepsilon} \tag{2.26}$$

Choose $\tau_{1,\varepsilon}, \tau_{2,\varepsilon}$ such that q_ε enters Γ_ε at time $t_1 - \tau_{1,\varepsilon}$ and leaves it at time $t_1 + \tau_{2,\varepsilon}$. Writing

$$\nabla V_{h,\varepsilon}(x) = \alpha(x)n + \beta(x)Px \tag{2.27}$$

for $n = n(q(t_1))$, P the projection onto the plane perpendicular to n and $\alpha, \beta \in \mathbb{R}$, we obtain that $|\beta(x)|$ is bounded on $r\varepsilon$-neighborhoods of $q(t_1)$ independently of ε for each $r \in \mathbb{R}$, whereas $\alpha(x)$ scales with ε^{-1}. Splitting the trajectories accordingly as

$$q_\varepsilon(t) = q_\varepsilon^{\parallel}(t) + q_\varepsilon^{\perp}(t) \quad \text{with} \quad q_\varepsilon^{\parallel} \in q(t_1) + \mathbb{R}M^{-1}n,\ q_\varepsilon^{\perp} \in M^{-1}P\mathbb{R}^d \tag{2.28}$$

we obtain that

$$\ddot{q}_\varepsilon^{\parallel}(t) = -M^{-1}\alpha(q_\varepsilon(t))n, \qquad \ddot{q}_\varepsilon^{\perp}(t) = -M^{-1}\beta(q_\varepsilon(t))Pq_\varepsilon(t) \tag{2.29}$$

It is not hard to see that $\tau_{2,\varepsilon} + \tau_{1,\varepsilon} = O(\varepsilon)$ and so $|\ddot{q}_\varepsilon^{\parallel}(t) + M^{-1}\alpha(q_\varepsilon^{\parallel}(t))n| \leq C$, $|\ddot{q}_\varepsilon^{\perp}(t)| \leq C$, on $(t_1 - \tau_{1,\varepsilon}, t_1 + \tau_{2,\varepsilon})$ for some suitable constant $C > 0$. This proves that in fact $\dot{q}_\varepsilon^{\perp}(t_1 - \tau_{2,\varepsilon}) - \dot{q}_\varepsilon^{\perp}(t_1 - \tau_{1,\varepsilon}) \to 0$ as $\varepsilon \to 0$ and, by energy conservation,

$$\dot{q}_\varepsilon(t_1 - \tau_{2,\varepsilon}) - \dot{q}_\varepsilon(t_1 - \tau_{1,\varepsilon}) \to \dot{q}^+ - \dot{q}^- \tag{2.30}$$

where $\dot{q}^+$ is given by (2.15) resp. (2.18) resp. (2.21). From here it is now straightforward to obtain that q_ε converges to $\tilde{q}$, where $\tilde{q}$ is linear on (t_0, t_1) and (t_1, t_2) and satisfies the jump condition (2.15) resp. (2.18) resp. (2.21) at t_1.

Since we do not wish to use the fact that q is explicitly given by these equations in our proof, we proceed as follows. As q_ε is a critical point of $I_{h,\varepsilon}$, for all $\varphi \in C_c^\infty((t_0, t_2), \mathbb{R}^d \times \mathbb{R}^d)$, the first variation

$$\delta I_{h,\varepsilon}(q, \varphi) = \int_{t_0}^{t_2} \dot{q}^T M \dot{\varphi} - \nabla V_{h,\varepsilon}(q)\varphi \, dt \tag{2.31}$$

vanishes. A standard approximation argument show that this remains true for all $\varphi \in W_0^{1,\infty}((t_0, t_2), \mathbb{R}^d \times \mathbb{R}^d)$. Now suppose that q' is another piecewise linear trajectory with $\tilde{q}(t_0) = q'(t_0)$, $\tilde{q}(t_2) = q'(t_2)$, which is nearby $\tilde{q}$ and meets the same energy jump surface Γ at time t_1'. Choose $\tau_{1,\varepsilon}', \tau_{2,\varepsilon}'$, such that q' enters, resp. leaves, Γ_ε at time $t_1' - \tau_{1,\varepsilon}'$ resp. $t_1' + \tau_{2,\varepsilon}'$

Now construct the following particular approximation q_ε' to q':

- $q_\varepsilon'(t)$ is linear on $(t_0, t_1' - \tau_{1,\varepsilon})$ and on $(t_1' + \tau_{2,\varepsilon}, t_2)$,
- $q_\varepsilon'(t_1' - \tau_{1,\varepsilon}) = q'(t_1' - \tau_{1,\varepsilon}')$ and $q_\varepsilon'(t_1' + \tau_{2,\varepsilon}) - q'(t_1' + \tau_{2,\varepsilon}') \to 0$ as $\varepsilon \to 0$ and
- $V_{h,\varepsilon}(q_\varepsilon'(t_1' + t)) = V_{h,\varepsilon}(q_\varepsilon(t_1 + t))$ for $t \in [-\tau_{1,\varepsilon}, \tau_{2,\varepsilon}]$.

It is not hard to see that, as $\varepsilon \to 0$, $q_\varepsilon' \to q'$ strongly in $W^{1,p}$ for all $p < \infty$ (and weak* in $W^{1,\infty}$), and in particular $I_{h,\varepsilon}(q_\varepsilon) \to I_h(\tilde{q})$, $I_{h,\varepsilon}(q_\varepsilon') \to I_h(q')$. In fact, due to the careful definition of q_ε', we also obtain control over the difference

$$I_{h,\varepsilon}(q_\varepsilon) - I_{h,\varepsilon}(q_\varepsilon') \to I_h(\tilde{q}) - I_h(q') \tag{2.32}$$

To see this, it suffices to note that $\displaystyle\int_{t_1' - \tau_{1,\varepsilon}}^{t_1' + \tau_{2,\varepsilon}} V(q_\varepsilon'(t))\, dt - \int_{t_1 - \tau_{1,\varepsilon}}^{t_1 + \tau_{2,\varepsilon}} V(q_\varepsilon(t))\, dt = 0$. In fact, since q_ε is a critical point of $I_{h,\varepsilon}$, we have

$$|I_{h,\varepsilon}(q_\varepsilon) - I_{h,\varepsilon}(q_\varepsilon')| = o(\|q_\varepsilon - q_\varepsilon'\|_{W^{1,2}}) \tag{2.33}$$

with a term $o(\cdot)$ independent of ε. (Consider a path $[0,1] \ni s \mapsto q_\varepsilon(\cdot, s)$ in the space of curves such that $q_\varepsilon(\cdot, 0) = q_\varepsilon$, $q_\varepsilon(\cdot, 1) = q'_\varepsilon$ and $V_{h,\varepsilon}(q_\varepsilon(t_1(s) + \cdot, s)) = V_{h,\varepsilon}(q_\varepsilon(t_1 + \cdot))$ on $[-\tau_{1,\varepsilon}, \tau_{2,\varepsilon}]$, where $t_1(s)$ interpolates between t_1 and t'_1, i.e., $t_1(0) = t_1$ and $t_1(1) = t'_1$.)

As a consequence we obtain

$$|I_h(\tilde{q}) - I_h(q')| = o(\|\tilde{q} - q'\|_{W^{1,2}}) \tag{2.34}$$

and thus that $\tilde{q}$ is a critical point of the action functional I_h on piecewise linear trajectories as identified in Section 2.2. Now since $q(0) = \tilde{q}(0)$ and $\dot{q}(0) = \dot{\tilde{q}}(0)$, we obtain $\tilde{q} = q$. Hence, q is the limit, as $\varepsilon \to 0$ of the smooth trajectories q_ε corresponding to the smooth Lagrangian $L_{h,\varepsilon}$. ■

2.2.3 Summary of the energy-stepping scheme

We close this section by summarizing the relations obtained in the foregoing and defining the *energy-stepping* approximation scheme resulting from a piecewise-constant approximation of the potential energy.

Definition 2.2.1 (Energy-stepping) Suppose $\left(t_k, q_k, \dot{q}_k^+\right)$ and a piecewise-constant approximation of the potential energy V_h are given. Let t_{k+1} and q_{k+1} be the time and point of exit of the rectilinear trajectory $q_k + (t - t_k)\dot{q}_k^+$ from the set $\{V = h\mathbb{Z}\}$. Let ΔV be the energy jump at q_{k+1} in the direction of advance. The updated velocity is, then,

$$\dot{q}_{k+1}^+ = \dot{q}_k^+ + \lambda_{k+1} M^{-1} n_{k+1} \tag{2.35}$$

where

$$\lambda_{k+1} = \begin{cases} -2\dfrac{\dot{q}_k^+ \cdot n_{k+1}}{n_{k+1}^T M^{-1} n_{k+1}}, \\ \qquad\qquad \text{if } \left(\dot{q}_k^+ \cdot n_{k+1}\right)^2 < 2\Delta V \left(n_{k+1}^T M^{-1} n_{k+1}\right), \\ \dfrac{-\dot{q}_k^+ \cdot n_{k+1} + \operatorname{sign}(\Delta V)\sqrt{\left(\dot{q}_k^+ \cdot n_{k+1}\right)^2 - 2\Delta V \left(n_{k+1}^T M^{-1} n_{k+1}\right)}}{n_{k+1}^T M^{-1} n_{k+1}}, \\ \qquad\qquad \text{otherwise.} \end{cases} \tag{2.36}$$

These relations define a discrete propagator

$$\Phi_h : \left(t_k, q_k, \dot{q}_k^+\right) \mapsto \left(t_{k+1}, q_{k+1}, \dot{q}_{k+1}^+\right) \tag{2.37}$$

that can be iterated to generate a discrete trajectory.

Algorithm 1 Energy-stepping integrator

Require: $V(q)$, q_0, $\dot{q}_0$, t_0, t_f and the energy step h

1: $k \leftarrow 0$
2: **while** $t_k < t_f$ **do**
3: $\quad t_{k+1} \leftarrow$ SMALLEST-ROOT$(V(q_k + (t_{k+1} - t_k)\dot{q}_k) - V(q_k) + \Delta V = 0)$
4: $\quad q_{k+1} \leftarrow q_k + (t_{k+1} - t_k)\dot{q}_k$
5: $\quad n_{k+1} \leftarrow \nabla V(q_{k+1})$
6: $\quad \dot{q}_{k+1} \leftarrow$ UPDATE-VELOCITIES$(\dot{q}_k, n_{k+1}, h)$
7: $\quad k \leftarrow k + 1$
8: **end while**

The implementation of the energy-stepping scheme is summarized in Algorithm 1. The algorithm consists of two methods. The first method SMALLEST-ROOT determines the earliest root of the equation

$$V(q_k + (t_{k+1} - t_k)\dot{q}_k) - V(q_k) + \Delta V = 0 \tag{2.38}$$

where ΔV can take values in $\{0, h, -h\}$. This task can be effectively accomplished by locating first, by means of an incremental search technique, a time interval containing the smallest positive root and then zeroing in on the root by means of an iterative procedure such as bisection, Newton-Raphson, or a combination thereof. The nature of the step, i.e., whether it consists of a diffraction by an

uphill energy step, a diffraction by a downhill energy step or a reflection at an uphill energy step, is identified simultaneously with the computation of t_{k+1}. The second method UPDATE-VELOCITIES is responsible for updating velocities according to Definition 2.2.1 and reduces to only two scenarios. The method is defined in Algorithm 2.

Algorithm 2 UPDATE-VELOCITIES$(\dot{q}_0, n_1, h)$

if $\dot{q}_0 \cdot n_1 \geq 0$ **then**
  $\Delta V \leftarrow h$ // Positive energy jump approached
else
  $\Delta V \leftarrow -h$ // Negative energy jump approached
end if
if $(\dot{q}_0 \cdot n_1)^2 \leq 2\Delta V \left(n_1^T M^{-1} n_1\right)$ **then**
  $\lambda \leftarrow -2\frac{\dot{q}_0 \cdot n_1}{n_1^T M^{-1} n_1}$ // Reflects back from energy barrier
else
  $\lambda \leftarrow \frac{-\dot{q}_0 \cdot n_1 + \text{sign}(\Delta V)\sqrt{(\dot{q}_0 \cdot n_1)^2 - 2\Delta V\left(n_1^T M^{-1} n_1\right)}}{n_1^T M^{-1} n_1}$ // Overcomes energy jump
end if
return $\dot{q}_0 + \lambda M^{-1} n_1$

It bears emphasis that energy-stepping requires the solution of no system of equations and, therefore, its complexity is comparable to that of explicit methods. However, the need to compute the root of a nonlinear function adds to the overhead of one application of the algorithm. It is still possible, however, that such overhead may be offset by the higher accuracy in particular applications. These and other trade-offs are investigated subsequently by way of numerical testing.

2.3 Conservation properties

Because the stepwise approximation of the potential energy (2.4) preserves all the symmetries of the system, and the discrete trajectories are exact trajectories of a Lagrangian system, energy-stepping exactly conserves all the momentum maps and the symplectic structure of the original Lagrangian system. These two types of conservation properties are examined next in turn.

2.3.1 Conservation of momentum maps

The theory of symmetry of smooth Lagrangian systems is well-known [57] but may stand a brief summary, specifically as it bears on the present application. We recall that a Lagrangian is a

function $L : TQ \to \mathbb{R}$, where Q is a smooth, or configuration, manifold of the system and TQ is the corresponding tangent bundle, consisting of pairs of configurations and velocities. For simplicity, we restrict attention to time-independent or autonomous Lagrangians. Let X denote some suitable topological space of trajectories $q : [0, T] \to Q$ joining fixed initial and final configurations $q(0)$ and $q(T)$, respectively. Then, the action integral $I : X \to \mathbb{R}$ over the time interval $[0, T]$ is

$$I(q) = \int_0^T L\big(q(t), \dot{q}(t)\big)\, dt \tag{2.39}$$

where we assume sufficient regularity of L and $q(t)$ for all mathematical operations to be well defined. According to Hamilton's principle, the physical trajectories of the system are the critical points of I, i.e., $q \in X$ is a trajectory if

$$\delta I(q, \varphi) = 0 \tag{2.40}$$

for all variations $\varphi \in C_c^\infty(0, T)$. Throughout this chapter the configurational space of interest is $Q = E(n)^N$, where $E(n)$ is the Euclidean space of dimension n (i.e., $d = nN$), and the Lagrangian is assumed to be of the form (2.1), with V smooth and bounded below, or (2.2) with V_h piecewise constant. An appropriate space of trajectories is $X = W^{1,p}([0, T])$, $p < \infty$, as demonstrated in Section 2.4. We note that in the case of a piecewise constant potential energy the Lagrangian is not differentiable and the Euler-Lagrange equations are not defined in the classical sense. However, the trajectories can still be understood as critical points of the action functional I_h as shown in Theorem 2.2.1.

Let G be a Lie group with Lie algebra $\mathfrak{g} = T_eG$. A left action of G on Q is a mapping $\Phi : G \times Q \to Q$ such that: i) $\Phi(e, \cdot) = \mathrm{id}$; ii) $\Phi(g, \Phi(h, \cdot)) = \Phi(gh, \cdot)$ $\forall g, h \in G$. Let $\xi \in \mathfrak{g}$. Then, the infinitesimal generator of Φ corresponding to ξ is the vector field $\xi_Q \in TQ$ given by

$$\xi_Q(q) = \frac{d}{dt}\left[\Phi(\exp(t\xi), q)\right]_{t=0} \tag{2.41}$$

The momentum map $J : TQ \to \mathfrak{g}^*$ defined by the action Φ then follows from the identity

$$\langle J(q,\dot{q}), \xi \rangle = \langle \partial_{\dot{q}} L(q,\dot{q}), \xi_Q(q) \rangle, \quad \forall \xi \in \mathfrak{g} \tag{2.42}$$

We say that the Lagrangian L is invariant under the action Φ if

$$L(\Phi_g(q), T\Phi_g(q)\dot{q}) = L(q,\dot{q}), \quad \forall g \in G,\ (q,\dot{q}) \in TQ \tag{2.43}$$

where we write $\Phi_g(\cdot) = \Phi(g,\cdot)$. Under these conditions, we additionally say that G is a symmetry group of the system and that Φ expresses a symmetry of the system. The classical theorem of Noether states that if L is invariant under the action Φ, then the corresponding momentum map J is a constant of the motion, i.e., it remains constant along trajectories. Classical examples include:

i) *Conservation of linear momentum.* In this case, $Q = E(n)^N$, $G = E(n)$ and $\Phi(u,q) = \{q_1+u, \ldots, q_N+u\}$ represents a rigid translation of the system by $u \in E(n)$. The corresponding momentum map is the total linear momentum of the system, $J = p_1 + \cdots + p_N$. If the Lagrangian is invariant under translations, then the total linear momentum is a constant of the motion.

ii) *Conservation of angular momentum.* In this case, $Q = E(n)^N$, $G = SO(n)$ and $\Phi(R,q) = \{Rq_1, \ldots, Rq_N\}$ represents a rigid rotation of the system by $R \in SO(n)$. The corresponding momentum map is the total angular momentum of the system, $J = q_1 \times p_1 + \cdots + q_N \times p_N$. If the Lagrangian is invariant under rotations, then the total angular momentum is a constant of the motion.

Conservation of energy can be fit into this framework by resorting to a space-time formulation in which time is regarded as a generalized coordinate, e. g., q_0. The corresponding space-time Lagrangian is $\mathbb{L}(\amalg, \amalg') = \mathbb{L}((q_0,q),(q_0',q')) = L(q, q'/q_0', q_0)$, where $L(q,\dot{q},t)$ is a general time-dependent Lagrangian. The space-time configuration manifold is $\mathbb{Q} = \mathbb{R} \times Q$. Let $G = \mathbb{R}$, $\Phi(s,q) = (q_0+s, q)$ is a time-shift and suppose that $\mathbb{L}$ is invariant under Φ, i.e., L is time-independent. Then $J =$

$L - \partial_{\dot{q}} L \cdot \dot{q} = -E$, i.e., the total energy of the system, is a constant of the motion.

A particularly appealing property of the piecewise constant approximation (2.4) of the potential energy is that it preserves all the symmetries of the system exactly. To verify this property, since the kinetic energy of the system is not approximated, it suffices to verify that V_h has all the symmetries of V. Thus, suppose that G is a symmetry group of V and Φ is an action that leaves V invariant, i.e.,

$$V \circ \Phi_g = V, \quad \forall g \in G \tag{2.44}$$

Then,

$$V_h \circ \Phi_g = (h\lfloor h^{-1} V \rfloor) \circ \Phi_g = h\lfloor h^{-1} V \circ \Phi_g \rfloor = h\lfloor h^{-1} V \rfloor = V_h \tag{2.45}$$

Similarly, we obtain for the smooth approximation of V_h introduced in (2.23)

$$V_{h,\varepsilon} \circ \Phi_g = \chi_\varepsilon \circ V \circ \Phi_g = \chi_\varepsilon \circ V = V_{h,\varepsilon} \tag{2.46}$$

It therefore follows from Noether's theorem that the corresponding momentum map J is constant along trajectories of the approximate Lagrangian $L_{h,\varepsilon}$. By Theorem 2.2.1 and a standard approximation argument then show that the momentum map J is constant along energy-stepping trajectories. It bears emphasis that what is conserved along the trajectories of the approximate Lagrangian L_h is the *exact*, time-continuous, momentum map of the original Lagrangian L. This is in contrast to discrete variational integrators, which conserve discrete form of the momentum maps, instead of the exact, time-continuous, momentum maps of the original Lagrangian L. Thus, in particular: if V is invariant under translations then energy-stepping conserves the total linear momentum $p_1 + \cdots + p_N$ of the system; and if V is invariant under rotations then energy-stepping conserves the total angular momentum $q_1 \times p_1 + \cdots + q_N \times p_N$ of the system.

As indicated above, the exact conservation of energy attendant to energy-stepping applied to time-independent Lagrangians follows in the manner described above from a space-time extension of the configuration manifold. However, exact energy conservation follows more directly as a conse-

quence of (2.12), i.e., energy conservation is built explicitly into the energy-stepping scheme.

2.3.2 Conservation of the symplectic structure

We now look at the Lagrangian systems defined by (2.1) and (2.2) from a Hamiltonian perspective. To this end we introduce the phase space $P = T^*Q$, consisting of pairs (q, p) of configurations $q \in Q = E(n)^N$ and momenta $p \in T_q^*Q$, and the Hamiltonian $H : P \to \mathbb{R}$ as

$$H(q,p) = \sup_{v \in T_qQ} \{p \cdot v - L(q,v)\} = \frac{1}{2} p^T M^{-1} p + V(q) \tag{2.47}$$

Likewise, the Hamiltonian corresponding to the approximate Lagrangian (2.2) is

$$H_h(q,p) = \sup_{v \in T_qQ} \{p \cdot v - L_h(q,v)\} = \frac{1}{2} p^T M^{-1} p + V_h(q) \tag{2.48}$$

We endow P with the canonical *symplectic two-form*

$$\Omega = dq_1 \wedge dp_1 + \cdots + dq_N \wedge dp_N \tag{2.49}$$

Then, the pair (P, Ω) defines a symplectic manifold.

We recall (cf. e. g., [1, 2, 32]) that a diffeomorphism $\varphi : P \to P$ is *symplectic* if it preserves the symplectic two-form, i.e., if

$$\Omega(T\varphi(z)\eta_1, T\varphi(z)\eta_2) = \Omega(\eta_1, \eta_2) \tag{2.50}$$

The symplecticity of Lipschitz maps has been investigated by Whitney [79] and by Gol'dshtein and Dubrovskiy [21] but their results do not apply to the present setting. However, it is possible to verify the symplecticity of energy-stepping directly. To this end, we may write $\varphi(q_0, p_0) = (q(t), p(t))$ and

$$T\varphi \equiv \begin{pmatrix} \partial_q \varphi_q & \partial_p \varphi_q \\ \partial_q \varphi_p & \partial_p \varphi_p \end{pmatrix} = \begin{pmatrix} Q_q & Q_p \\ P_q & P_p \end{pmatrix} \tag{2.51}$$

then φ is symplectic if

$$P_p^T Q_p = Q_p^T P_p \tag{2.52}$$

$$P_p^T Q_q = Q_p^T P_q + I \tag{2.53}$$

$$Q_q^T P_q = P_q^T Q_q \tag{2.54}$$

As shown in the appendix, these relations are identically satisfied by energy-stepping, which establishes the symplecticity of the scheme.

2.3.3 Summary of the conservation properties of the energy-stepping scheme

The results proven in the foregoing are collected and summarized in the following theorem.

Theorem 2.3.1 *The energy-stepping time-integration scheme is a symplectic-energy-momentum time-reversible integrator with automatic selection of the time step size. In particular, the scheme conserves exactly all the momentum maps of the original Lagrangian.*

Symmetry or time-reversibility of energy-stepping follows directly from the definition of the scheme. The automatic time-step selection property also follows by construction. In particular, in regions where the potential energy gradient ∇V is steep, the energy jumps are more closely spaced and the resulting time steps are small. By contrast, if the potential energy gradient is small, the resulting time steps are comparatively large. We again emphasize that the momentum maps conserved by energy-stepping are precisely those conserved by the original system. This is in contrast to variational integrators, which conserve discrete momentum maps that differ from the momentum maps of the original system in general.

2.4 Convergence analysis

Our aim in this section is to ascertain conditions under which a sequence q_h of energy-stepping trajectories with potential V_h and initial conditions $q_h(0) = q_0$ and $\dot{q}_h(0^+) = \dot{q}_0$ converges in a suitable sense to a solution q of the original equations of motion as $h \to 0$. For the sake of simplicity, we consider the motion $t \mapsto q(t) \in \mathbb{R}^d$ of a particle in a smooth potential energy landscape V with diagonal mass matrix of the form $M = m\,\mathrm{Id}$. The convergence analysis presented in the following carries over unchanged to the general case, albeit at the expense of slightly more cumbersome notation. The trajectories q are stationary points of the action functional

$$I(q) = \int_0^T \frac{m}{2}(\dot{q}(t))^2 - V(q(t))dt \tag{2.55}$$

For smooth trajectories, this principle is equivalent to the satisfaction of the equations of motion

$$m\ddot{q}(t) = -\nabla V(q(t)), \quad \text{given} \quad q(0) = q_0, \quad \dot{q}(0) = \dot{q}_0 \tag{2.56}$$

We will assume throughout that V is bounded from below and, without loss of generality, we may specifically assume that $V \geq 0$.

Within this setting, the aim of convergence analysis is to elucidate the limiting behavior of the stationary points of the approximating functionals I_h that arise when the smooth potential V is approximated by piecewise constant potentials V_h. However, due to the non-differentiability of V_h, the first variation

$$\delta I_h(q, \varphi) := \lim_{s \to 0} \frac{1}{s}\Big(\int_0^T \frac{m}{2}(\dot{q} + s\dot{\varphi})^2 - V_h(q + s\varphi)dt - \int_0^T \frac{m}{2}(\dot{q})^2 - V_h(q)dt \Big) \tag{2.57}$$

does not exist in general for arbitrary $\varphi \in C_c^\infty(0, t; \mathbb{R}^d)$. The equation of motion (2.56) needs to be adapted to this situation by allowing the approximating acceleration to be a measure, specifically, a sum of suitably rescaled Dirac masses. With this extension, we may summarize the the energy-stepping scheme formulated in Section 2.2.3 by saying that the approximating sequences satisfy the

Euler-Lagrange equations

$$m\ddot{q} = \sum_{t \in S(q)} \alpha(q(t), \dot{q}(t^-), \dot{q}(t^+)) n(q(t)) \delta_t \tag{2.58}$$

in the sense of $\mathbb{R}^d$-valued measures, where

$$\alpha(q(t), \dot{q}(t^-), \dot{q}(t^+)) = \begin{cases} -2m\dot{q}(t^-) \cdot n(q(t)) & \text{if } \sqrt{\frac{2h}{m}} > \dot{q}(t^-) \cdot n(q(t)) > 0 \\ \dfrac{-h}{\left| \frac{\dot{q}(t^+) + \dot{q}(t^-)}{2} \cdot n(q(t)) \right|} & \text{otherwise} \end{cases} \tag{2.59}$$

$n(q) = \frac{\nabla V(q)}{|\nabla V(q)|}$ and $S(q)$ denotes the set of times t where $\dot{q}(t)$ jumps. Indeed, as $\dot{q}$ is only piecewise constant, $\ddot{q}$ is only defined as a distribution. In order to identify $\ddot{q}$, let $\varphi : (0, \infty) \to \mathbb{R}^d$ be a smooth, compactly supported function. We have to verify that

$$-\int_0^\infty m\dot{q}(s) \cdot \dot{\varphi}(s)\, ds = \int_0^\infty \varphi(s) \cdot\, d\left(\sum_{t \in S(q)} \alpha(q(t), \dot{q}(t^-), \dot{q}(t^+)) n(q(t)) \delta_t \right)(s) = \sum_{t \in S(q)} \alpha(q(t), \dot{q}(t^-), \dot{q}(t^+)) n(q(t)) \cdot \varphi(t) \tag{2.60}$$

In fact, by recourse to a partition of unity it suffices to verify this identity for φ supported on intervals (t', t'') that contain at most one point of $S(q)$. If $S(q) \cap (t', t'') = \emptyset$, then $\dot{q}$ is constant on (t', t'') and both sides of (2.60) are zero. Now let $S(q) \cap (t', t'') = \{t\}$ and denote the velocity $\dot{q}$ prior to t, resp. after t, by $\dot{q}(t^-)$, resp. $\dot{q}(t^+)$. Then (2.60) is equivalent to

$$-\int_{t'}^{t''} m\dot{q}(s) \cdot \dot{\varphi}(s)\, ds = \alpha(q(t), \dot{q}(t^-), \dot{q}(t^+)) n(q(t)) \cdot \varphi(t) \tag{2.61}$$

Splitting the integral as $\int_{t'}^{t} + \int_{t}^{t''}$ we obtain

$$-\int_{t'}^{t''} m\dot{q}(s) \cdot \dot{\varphi}(s)\, ds = m(\dot{q}(t^+) - \dot{q}(t^-)) \cdot \varphi(t) \tag{2.62}$$

and Equation (2.60) follows with

$$\alpha n = m(\dot{q}(t^+) - \dot{q}(t^-)) \tag{2.63}$$

This identity in turn is consistent with the more direct approach followed in Section 2.2.3. Indeed, if the trajectory reflects at an energy level surface, (2.63) is directly implied by Definition 2.2.1. If, contrariwise, an energy level surface is crossed Definition 2.2.1 gives

$$m\left(\dot{q}(t^+) - \dot{q}(t^-)\right) = m\left(-\dot{q}(t^-)\cdot n + \operatorname{sign}(\Delta V)\sqrt{(\dot{q}(t^-)\cdot n)^2 - 2\Delta V m^{-1}}\right) n =: \lambda n \tag{2.64}$$

with $\Delta V = V_h(q(t^+)) - V_h(q(t^-))$, $h = |\Delta V|$. Multiplying with $\dot{q}(t^-) + \dot{q}(t^+)$, we obtain from Equation (2.12)

$$-2\Delta V = m\left((\dot{q}(t^+))^2 - (\dot{q}(t^-))^2\right) = \lambda\left(\dot{q}(t^-) + \dot{q}(t^+)\right)\cdot n \tag{2.65}$$

whence indeed

$$\lambda = \frac{-\Delta V}{\frac{\dot{q}(t^+) + \dot{q}(t^-)}{2}\cdot n} = \frac{-h}{\left|\frac{\dot{q}(t^+) + \dot{q}(t^-)}{2}\cdot n\right|} = \alpha \tag{2.66}$$

In general, the approximating sequences thus defined may fail to converge to trajectories of the original system as $h \to 0$. The nature of the difficulty is illustrated by the following three examples.

2.4.1 Failure of convergence

The first example shows that convergence does not necessarily hold true even for one-dimensional problems.

Example 2.4.1 *For definiteness suppose that $m = 1$, $q_0 = 0, \dot{q}_0 = 1$ and V is given by $V(x) = \frac{1}{2}x^2$. Then* (2.56) *describes a harmonic oscillator and the trajectory is $q(t) = \sin(t)$. Similarly, q_h increases monotonically as long as $V_h(q_h) < \frac{1}{2}$. Let $k = \lfloor \frac{1}{2h} \rfloor$. Then on $\{t : q_h(t) \in (\sqrt{2kh}, \sqrt{2(k+1)h})\}$ the approximating velocity is $\dot{q}_h = \sqrt{1 - 2kh}$, and the time from reaching the top energy level kh until*

bouncing at $q_h = \sqrt{2(k+1)h}$ *is*

$$\Delta t = \frac{\sqrt{2(k+1)h} - \sqrt{2kh}}{\sqrt{1-2kh}} \approx \frac{h}{\sqrt{\frac{1}{2} - kh}} \tag{2.67}$$

In particular, q_h *does not converge to* q *if the last expression does not converge to zero as* $h \to 0$.

This example shows that there may be sequences h_i such that the stationary trajectories of I_{h_i} do not converge to stationary trajectories of I. Of course, this does not rule out the existence of sequences h_i for which convergence does take place, and non-convergence situations of the type illustrated by the example can be readily avoided by a proper choice of sequence. The second example shows that convergence of q_h to q may fail due to *non-transversality* of the velocity with respect to ∇V.

Example 2.4.2 *Let* $V : \mathbb{R}^2 \to \mathbb{R}_+$ *and suppose that* $V(x,y) = \frac{1}{2}x^2$. *If* $m = 1$, $q_0 = (1,0)^T, \dot{q}_0 = (0,1)^T$, *then the solution* q *of* (2.56) *is given by* $q(t) = (\cos(t), t)^T$. *However, the solutions* q_h *of* (2.58) *are uniform motions in* y*-direction, i.e.,* $q_h(t) = (0,t)^T$, *independently of* h.

While in this example the lack of convergence is due to the approximating trajectories not *feeling* the force in x-direction, we show in the last example that even if there are no long time intervals over which none of the level sets $\{V = kh\}$ is hit, the approximating trajectory may fail to converge to a solution of (2.56).

Example 2.4.3 *Let* $V : \mathbb{R}^2 \to \mathbb{R}_+$, $V(x,y) = \frac{1}{2}(x^2+y^2)$ *and assume that* $m = 1$, $q_0 = (1,0)^T$ *and* $\dot{q}_0 = (0,\gamma)^T$, $\gamma > 0$. *Let* $k = \lfloor \frac{1}{2h} \rfloor$. *Elementary geometric considerations show that* $\operatorname{dist}(q_0, \{V = (k+1)h\}) \leq h + o(h)$ *and, if* q_h *hits the set* $\{V = (k+1)h\}$, *then* $\gamma\sqrt{2h} + o(\sqrt{h}) > \dot{q}_h(t-) \cdot n(q(t)) > 0$. *If* $\gamma < 1$, *then* q_h *bounces and then indeed* q_h *bounces whenever it hits the set* $\{V = (k+1)h\}$, *that is the trajectory describes a polygon approximating the circle with radius* 1 *centered at* 0. *However, the solution* q *of the continuum problem is* $q(t) = (\cos(t), \gamma\sin(t))^T$ *and, hence, for* $\gamma < 1$, q_h *can not converge to* q *(see Figure 2.3 for an illustration).*

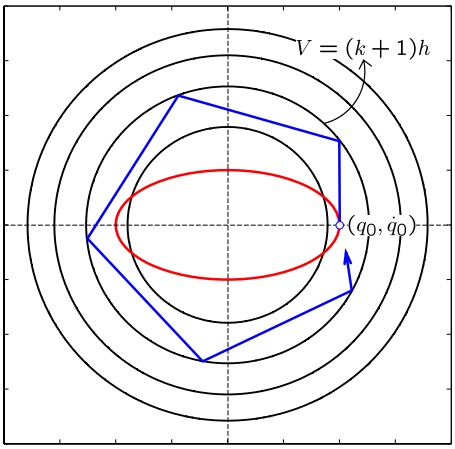


Figure 2.3: Illustration of Example 2.4.3. Red line: Exact solution. Blue line: Energy-stepping solution.

These examples suggest that the lack of convergence is connected to very slow motion in the direction parallel to ∇V. As observed in Section 2.4.3, in one dimension this can be overcome by a suitable selection of the initial conditions. This is due to the fact that in one dimension, energy conservation furnishes a one-to-one correspondence between the initial velocity and the velocity at a later time near a turning point. Analogously in higher dimensions, a proper choice of the initial conditions could remedy the lack of the convergence that might appear at the first point the continuum trajectory becomes non-transversal to an energy jump surface. However, there seems to be no natural condition that would guarantee convergence also at later turning points. (Note that energy conservation only gives control over the absolute value $|\dot{q}|$. In higher dimension $|\dot{q} \cdot \nabla V|$ can be very small even if $|\dot{q}|$ is large.) Examples 2.4.2 and 2.4.3 have to be understood in this vein: Even if the trajectory is well behaved initially on the time interval $[0, t_0]$, for example, convergence might fail after time $t_0 > 0$. This is indeed the case if $(q(t_0), \dot{q}(t_0))$ assumes the value denoted $(q_0, \dot{q}_0)$ in Example 2.4.2 or 2.4.3.

2.4.2 Convergence of transversal trajectories

In view of the preceding examples we introduce the following definition.

Definition 2.4.1 *Let q_h be a sequence of solutions of* (2.58) *and set $\Delta q_{h,k} := q_h(t_{k+1}) - q_h(t_k)$,*

where $S(q_h) = \{t_1, \dots, t_{N_h}\}$ is the jump set of $\dot{q}_h(t)$. We say that q_h is transversal if

$$\lim_{h\to 0} \min_k \frac{|n(t_k)\cdot \dot{q}_h(t_k^-)|}{\sqrt{h}} = \infty \quad \textit{and} \quad \lim_{h\to 0} \max_k \frac{|\Delta q_{h,k}|}{\sqrt{h}} = 0$$

Then, we have the following convergence result.

Theorem 2.4.4 *Let q_0 and $\dot{q}_0$ be given such that $\nabla V(q_0)\cdot \dot{q}_0 \neq 0$ and suppose q_h solves (2.58) with $q_h(0) = q_0$ and $\dot{q}_h(0^+) = \dot{q}_0$ for $h > 0$. Then if the sequence q_h is transversal, $q_h \to q$ in $W^{1,p}(0,T)$ for each fixed $T > 0$ and all $p < \infty$, where q is the unique solution of (2.56).*

Proof. Since V is bounded from below, we have $\|q_h\|_{L^\infty}, \|\dot{q}_h\|_{L^\infty} \leq C$ for some constant C independent of h. Therefore, by passing if necessary to a subsequence we can assume that

$$q_h \overset{*}{\rightharpoonup} q \quad \text{in } W^{1,\infty} \tag{2.68}$$

for some $q \in W^{1,\infty}$, and thus $q_h \to q$ uniformly. Let $\varphi \in C([0,T];\mathbb{R}^d)$. Then

$$\int_0^T h \sum_{t\in S(q_h)} \alpha(q_h(\tau), \dot{q}_h(\tau^-), \dot{q}_h(\tau^+)) n(q_h(\tau))\cdot\varphi(\tau)\delta_t(d\tau) = h\sum_{k=1}^{N_h} \alpha(q_h(t_k), \dot{q}_h(t_k^-), \dot{q}_h(t_k^+)) n(q_h(t_k))\cdot\varphi(t_k) \tag{2.69}$$

where $t_1 < t_2 < \ldots < t_N$ are such that $\{t_1,\dots,t_N\} = S(q_h) = \{t : V(q_h(t)) \in h\mathbb{N}\}$. By transversality, $|\Delta q_{h,k}| \to 0$ as $h \to 0$ uniformly in k and $\dot{q}_h(t_k^+)\cdot n(q_h(t_{k+1})) = \dot{q}_h(t_{k+1}^-)\cdot n(q_h(t_{k+1})) \notin (0, \sqrt{\frac{2h}{m}})$. Therefore, we have

$$h = |V(q_h(t_{k+1})) - V(q_h(t_k))| = \left|\nabla V(q_h(t_k))\cdot\Delta q_{h,k} + \frac{1}{2}D^2V(\lambda_k q_h(t_k) + (1-\lambda_k)(t_{k+1}))(\Delta q_{h,k})^2\right| = \left|\nabla V(q_h(t_k))\cdot\dot{q}_h(t_k+)(t_{k+1}-t_k) + \frac{1}{2}D^2V(q_h(\tau_k))(\Delta q_{h,k})^2\right| \tag{2.70}$$

for some $\lambda_k \in (0,1)$, resp., $\tau_k \in (t_k, t_{k+1})$. Again, by transversality we obtain

$$h = |\nabla V(q_h(t_k)) \cdot \dot{q}_h(t_k+)|(t_{k+1} - t_k) + o(h) \tag{2.71}$$

Next we note that $|\alpha| \leq C\sqrt{h}$ since, if $\dot{q}(t^-) \cdot n(q(t)) \geq \sqrt{\frac{2h}{m}}$, then

$$\frac{\dot{q}(t^+) + \dot{q}(t^-)}{2} \cdot n(q(t)) \leq \frac{\dot{q}(t^-)}{2} \cdot n(q(t)) \tag{2.72}$$

whereas, if $\dot{q}(t^-) \cdot n(q(t)) \leq 0$, then

$$\frac{\dot{q}(t^+) + \dot{q}(t^-)}{2} \cdot n(q(t)) \leq \frac{\dot{q}(t^+)}{2} \cdot n(q(t)) \leq -C\sqrt{h} \tag{2.73}$$

From these bounds we obtain $|\dot{q}_h(t_k^-) - \dot{q}_h(t_k^+)| \leq C\sqrt{h}$. As $|n(t_k) \cdot \dot{q}_h(t_k^-)| \gg \sqrt{h}$, this in turn implies that

$$h = \left| \nabla V(q_h(t_k)) \cdot \frac{\dot{q}_h(t_k^-) + \dot{q}_h(t_k^+)}{2} \right| (t_{k+1} - t_k) + o(h) \tag{2.74}$$

It now follows that

$$\int_0^T h \sum_{t \in S(q_h)} \frac{\nabla V(q_h(\tau)) \cdot \varphi(\tau)}{|\nabla V(q_h(\tau)) \cdot \frac{\dot{q}_h(\tau^+) + \dot{q}_h(\tau^-)}{2}|} \delta_t(d\tau) = h \sum_{k=1}^{N_h} \frac{\nabla V(q_h(t_k)) \cdot \varphi(t_k)}{|\nabla V(q_h(t_k)) \cdot \frac{\dot{q}_h(t_k^+) + \dot{q}_h(t_k^-)}{2}|} = \sum_{k=1}^{N_h} \nabla V(q_h(t_k)) \cdot \varphi(t_k)(t_{k+1} - t_k) + o(1) \tag{2.75}$$

where we note that $N_h \leq Ch^{-1}$. As $q_h \to q$ uniformly as $h \to 0$ and

$$\max_k |t_{k+1} - t_k| \leq \frac{|\Delta q_{h,k}|}{|\dot{q}_h(t_{k+1}^-)|} \ll \frac{\sqrt{h}}{\sqrt{h}} = 1 \tag{2.76}$$

we obtain by a Riemann sum argument that the last term converges to $\displaystyle\int_0^T \nabla V(q(t)) \cdot \varphi(t) dt$. We have thus shown that the right hand side of (2.58) converges weak-* in measure to $-\nabla V(q)$. But then, as $\ddot{q}_h \to \ddot{q}$ in the sense of distributions, we can take the limit $h \to 0$ on both sides of the

Euler-Lagrange equations (2.58) to conclude that q satisfies the limiting equation

$$m\ddot{q}(t) = -\nabla V(q(t)) \tag{2.77}$$

We next have to show that q satisfies the proper initial conditions. Clearly $q(0) = q_0$ since $q_h \to q$ uniformly. To prove $\dot{q}(0) = \dot{q}_0$, we may assume, without loss of generality, that $\dot{q}_0 \cdot n(x) \geq c$ for some $c > 0$ and all x in a neighborhood $\mathcal{U}$ of q_0 (the case $\leq -c$ is analogous). As the velocities are bounded in terms of V and $\dot{q}_0$, we may choose $s > 0$ such that $q_h(t) \in \mathcal{U}$ for all $h > 0$ and all $t \in [0, s]$. Now let $s_h \in S(q_h)$ be the smallest time in $[0, s]$ such that $\dot{q}_h(s_h^+) \cdot n(q(s_h)) \leq \frac{c}{2}$ if such a time exists, $s_h := s$ otherwise. Then, since $\#(S(q_h) \cap [0, s_h]) \leq C s_h h^{-1}$ and $\alpha \leq Ch$ on $q_h|_{(0,s)}$, it follows that

$$|\dot{q}_h(t) - \dot{q}_h(\tilde{t})| \leq C s_h \quad \text{for} \quad \tilde{t} \in [0, s_h] \tag{2.78}$$

In particular, if $s_h < s$, then $|\dot{q}_0 - \dot{q}_h(s_h^+)| \geq \frac{c}{2}$ and therefore $s_h \geq C$ for some constant $C > 0$. Choosing s sufficiently small, again from (2.78) we deduce that

$$|\dot{q}_h(0^+) - \dot{q}_h(t)| \leq Cs \quad \text{for} \quad t \in [0, s] \tag{2.79}$$

and, since $\dot{q}_h|_{[0,s]} \rightharpoonup \dot{q}|_{[0,s]}$ in $L^\infty(0, T)$,

$$|\dot{q}_h(0^+) - \dot{q}(t)| \leq Cs \quad \text{for} \quad t \in [0, s] \tag{2.80}$$

Now sending $s \to 0$ proves the claim. By the uniqueness of the solutions of (2.56), we additionally conclude that the entire sequence q_h converges weak-* to q in $W^{1,\infty}$. But then $q_h \to q$ uniformly and, hence, $\|\dot{q}_h\|_{L^2}^2 \to \|\dot{q}\|_{L^2}^2$ since

$$\frac{1}{2}(\dot{q}_h)^2 + V(q_h) = E_h \to E = \frac{1}{2}(\dot{q})^2 + V(q) \tag{2.81}$$

By virtue of the boundedness of $|\dot{q}_h|$, the convergence is indeed strong in $W^{1,p}$ for all $p < \infty$. ∎

2.4.3 Discussion of the transversality condition

We conclude this section with a formal argument that, for generic initial data, we may expect convergence of the approximating trajectories beyond turning points. This conjecture is born out by the numerical tests presented in subsequent sections, which suggest that the energy-stepping trajectories may be expected to converge in practice.

In one dimension the argument can be made precise. Assume that $\dot{q}_h(0) \neq 0$. Then, the conserved total energy of the system is $E = \frac{m}{2}\dot{q}_0^2 + V_h(q_0)$ and the exact trajectory is a periodic motion between two turning points. We additionally suppose that $V' \neq 0$ at the turning points. Within a finite time interval these points are met finitely often. Under these conditions, convergence fails if and only if the time spent by the trajectory on the highest energy level near the turning points does not converge to zero. As in Example 2.4.1, we see that such is the case if and only if

$$\limsup_{h\to 0} \Delta t = \limsup_{h\to 0} \frac{\Delta q}{\sqrt{2\frac{E-kh}{m}}} \neq 0 \tag{2.82}$$

with $k = \left\lfloor \frac{E}{h} \right\rfloor$ and $\Delta q = q_{k+1} - q_k$ with $V(q_k) = kh$ and $V(q_{k+1}) = kh + h$. Now for Δq near a turning point q we obtain

$$\Delta q = \frac{h}{V'(q)} + O(h^2) \tag{2.83}$$

and Δq scales with h. From (2.82), it follows that convergence fails if

$$\liminf_{h\to 0} \frac{E - kh}{h^2} < \infty \tag{2.84}$$

However, generically $E - kh$ is of the order h. The probability of spending a finite amount of time near a turning point is therefore asymptotically less than $h^\alpha \ll 1$ for any $\alpha < 1$. Thus, as surmised in connection with Example 2.4.1, it is always possible to select sequences h_i such that the stationary trajectories of I_{h_i} converge to trajectories of I. Figure 2.4 illustrates the typical behavior of $\Delta t(h)$ when the trajectory is in the vicinity of a turning point. It is evident in the figure that $\Delta t(h)$ indeed becomes vanishingly small as $h \to 0$ on the complement of an exceptional set of small measure.

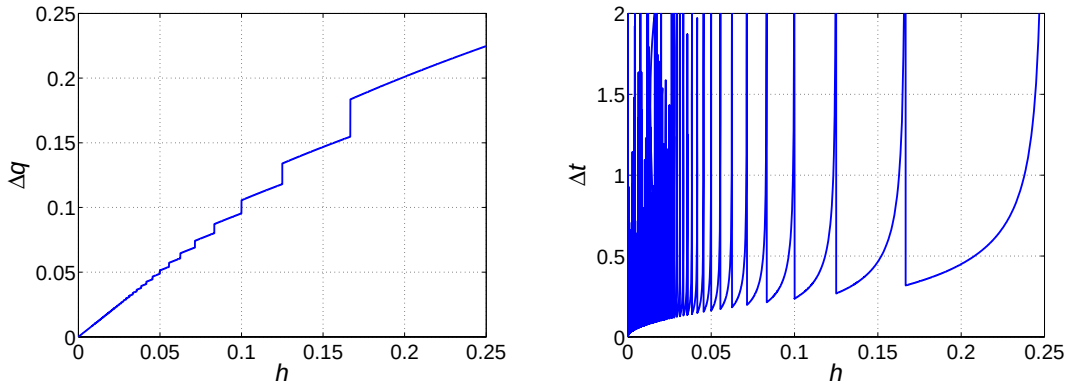


Figure 2.4: The behavior of $\Delta t(h)$ near a turning point.

In arbitrary dimension, for smooth V the motion near a turning point q may be decomposed into a component parallel to the energy jump surface and a transverse component in the direction of $\nabla V(q)$ (cf. Equation (2.28)). In this transverse direction the motion is ostensibly one-dimensional and we expect the preceding argument applies. In order to render this formal argument rigorous we would need an estimate ensuring that the energy associated with the transverse motion at the turning point is not too small. However, no such estimate is known to us for general systems.

2.5 Numerical examples

Next, we present selected examples of application that showcase the conservation, accuracy, long-term behavior and convergence properties of energy-stepping. We select two areas of application where those properties play an important role. For instance, good long-term behavior is essential for purposes of computing equilibrium thermodynamic properties in molecular dynamics. Good energy conservation properties are likewise important in elastic collision problems, especially in many-body problems where the fine structure of collisions may influence significantly the overall behavior of the system. We illustrate the performance of energy-stepping in those areas of application by means of three examples: the dynamics of a frozen argon cluster, the dynamics of a spinning elastic cube, and the dynamics of two colliding elastic spheres.

2.5.1 Frozen argon cluster

Molecular dynamics falls squarely within the framework considered in this chapter. Many applications in materials science, such as the calculation of free energies, require the integration of the system over long periods of time. In these applications, it is essential that the time integrator have good long-time behavior, such as conferred by symplecticity and exact conservation properties.

The velocity Verlet scheme, which is identical to Newmark's method (cf., for example [42]) of structural dynamics, is perhaps the most widely used time-integration scheme in molecular dynamics. The constant time-step velocity Verlet scheme is symplectic-momentum preserving and, therefore, does not conserve energy. As it is often the case with symplectic-momentum preserving methods, it nevertheless has good energy-conservation properties for sufficient small time steps. However, due to the conditional stability of the method the time-step is constrained by the period of thermal vibrations of the atoms, which renders calculations of equilibrium thermodynamic properties exceedingly costly. The development of integration schemes that alleviate or entirely eliminate the time-step restrictions of explicit integration in molecular dynamics applications is the subject of ongoing research (cf., for example [36, 38]).

We proceed to illustrate the performance of energy-stepping in molecular dynamics applications by analyzing the dynamics of a simple argon cluster. Specifically, we consider the numerical experiment proposed by Biesiadecki and Skeel [5]. The experiment concerns the two-dimensional simulation of a seven-atom argon cluster, six atoms of which are arranged symmetrically around the remaining central atom, Figure 2.5. The atoms interact via the pairwise Lennard-Jones potential

$$\phi(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] \tag{2.85}$$

where r is the distance between two atomic centers, $\varepsilon/k_B = 119.8$ K and $\sigma = 0.341$ nm are material constants for Argon, and $k_B = 1.380658 \cdot 10^{-23}$ J/K is the Boltzmann's constant. In addition, the mass of an argon atom is $m = 66.34 \cdot 10^{-27}$ kg. The initial positions of the atoms are slightly perturbed about the configuration that minimizes the potential energy of the cluster. The initial

velocities are chosen such that the total linear momentum is zero and the center of mass of the cluster remains fixed. The corresponding total energy of the cluster is $E/\varepsilon = -10.519$. Table 2.1 summarizes the initial conditions for the simulation.

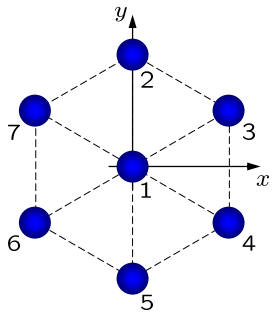


Figure 2.5: Frozen argon cluster.

Atom	1	2	3	4	5	6	7
Position [nm]	0.00	0.02	0.34	0.36	−0.02	−0.35	−0.31
	0.00	0.39	0.17	−0.21	−0.40	−0.16	0.21
Velocity [nm/nsec]	−30	50	−70	90	80	−40	−80
	−20	−90	−60	40	90	100	−60

Table 2.1: Frozen argon cluster: initial conditions.

Three different energy steps are employed in the energy-stepping calculations: $h_1 = E_0/100$, $h_2 = E_0/60$ and $h_3 = E_0/30$. The time steps employed in the velocity Verlet calculations are $\Delta t_1 = 56.98$ fsec, $\Delta t_2 = 87.56$ fsec and $\Delta t_3 = 124.88$ fsec. These time steps correspond to the average time steps resulting from the respective energy-stepping calculations. The total duration of the analysis is 100 nsec.

Figure 2.6 shows the evolution in time of the total energy. As expected, energy-stepping is energy preserving regardless the energy-step employed. By way of sharp contrast, whereas velocity Verlet has remarkable energy behavior for the short time step, it becomes unstable for the intermediate time step and blows up for the larger time step. This blow-up behavior is expected owing to the conditional stability of velocity Verlet.

In order to assess the long-time behavior of energy-stepping, we study the qualitative behavior of the trajectories of the argon atoms for a time window of $[99.95, 100]$ nsec, corresponding to the

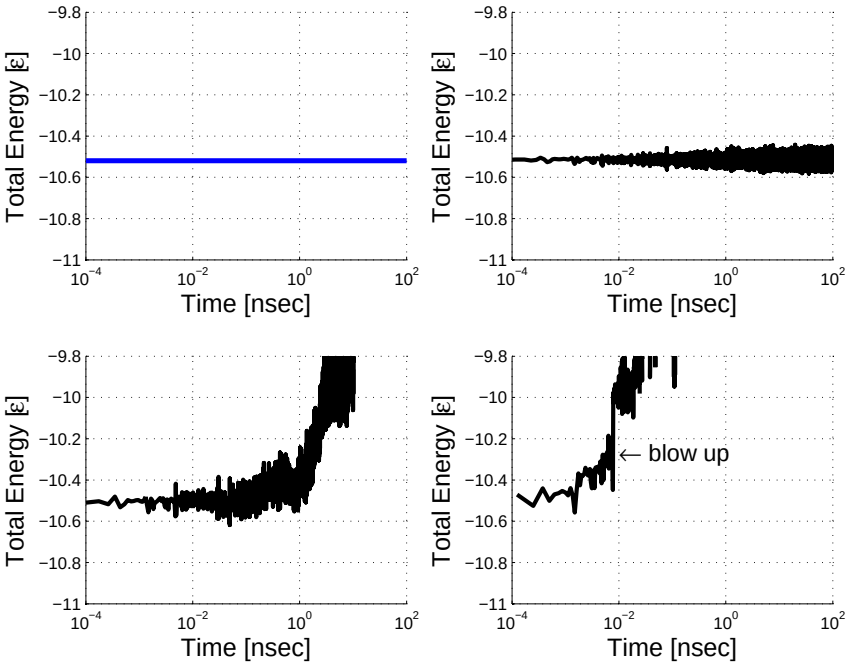


Figure 2.6: Frozen argon cluster. Top left: Energy-stepping. Top-right: Velocity Verlet, $\Delta t_1 = 56.98$ fsec. Bottom-left: Velocity Verlet, $\Delta t_2 = 87.56$ fsec. Bottom-right: Velocity Verlet, $\Delta t_3 = 124.88$ fsec.

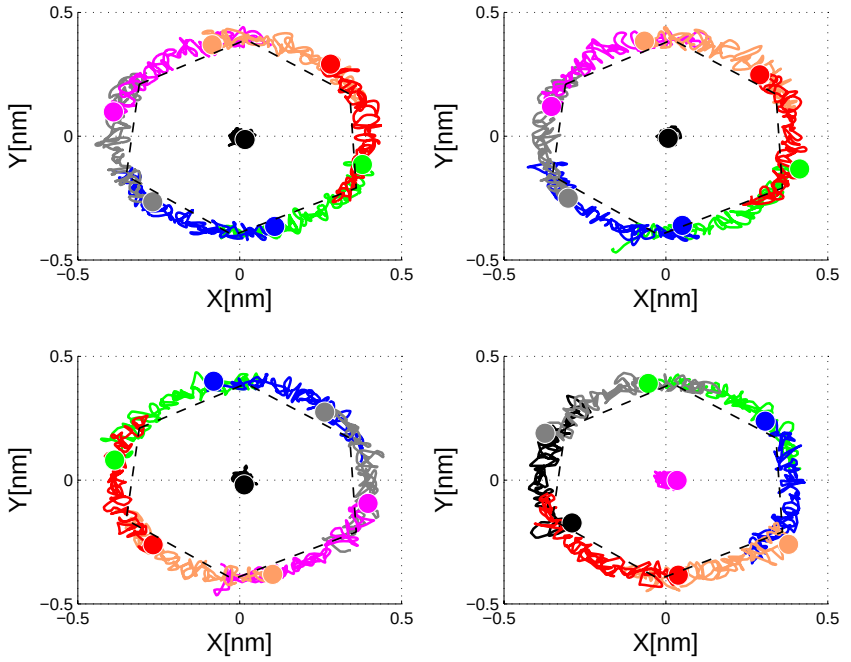


Figure 2.7: Trajectories of argon atoms for a time window of $[99.95, 100]$ nsec. The dashed line hexagon represents the initial position of the atoms. Top-left: velocity Verlet, $\Delta t = 10$ fsec. Top-right: Energy-stepping, $h_1 = E_0/100$. Bottom-left: Energy-stepping, $h_2 = E_0/60$. Bottom-right: Energy-stepping, $h_3 = E_0/30$.

last $50,000$ fsec of the simulations. A velocity Verlet solution computed with a time step equal to $\Delta t = 10$ fsec is presumed to be ostensibly converged and is used by way of reference. Trajectories of the seven argon atoms, each represented by a different color, are depicted in Figure 2.7 and the configuration at $t = 0$ is represented by the dashed line hexagon. A notable feature of the energy-stepping trajectories is that, even for large time steps, they remain stable over long periods of time.

Finally, the convergence statement of Theorem 2.4.4 is illustrated by Figure 2.8. We first recall that, as demonstrated in Section 2.4, the approximate space of trajectories is $X = W^{1,p}(0,T)$, $p < \infty$, and of particular interest is the space $H^1(0,T) := W^{1,2}(0,T)$. Then, distances are measured with respect to the H^1-norm

$$\|q\|_{H^1(0,T)} = \left(\int_0^T \left(|q(t)|^2 + |\dot{q}(t)|^2 \right) dt \right)^{1/2} \tag{2.86}$$

and a relative H^1-error is defined by

$$H^1\text{-error} = \frac{\left| \|q_h\|_{H^1(0,T)} - \|q\|_{H^1(0,T)} \right|}{\|q\|_{H^1(0,T)}} \tag{2.87}$$

where $\|q\|_{H^1(0,T)}$ is estimated from numerical results. The relative H^1-error is shown on the left side of Figure 2.8 as a function of the energy step, for $T = 1$ nsec. The slope of the convergence plot is also directly related to the rate of convergence in energy step h, that is H^1-error $= O(h^r)$. This gives an estimated rate of convergence in the H^1-norm of $r \simeq 1/2$. Furthermore, it is interesting to observe on the right side of Figure 2.8 that the average and the maximum time steps selected by energy-stepping are $O(h)$ and $O(h^{1/2})$, respectively. Therefore, the sequence q_h is indeed convergent, transversal, and $\lim_{h\to 0} \max_k \Delta t_k = 0$, as expected from the discussion in Section 2.4.3.

2.5.2 Finite-element models

Next we consider finite-dimensional Lagrangian systems obtained by a finite-element discretization of the action of a nonlinear elastic solid (cf., e. g., [51] for details of finite-element approximation

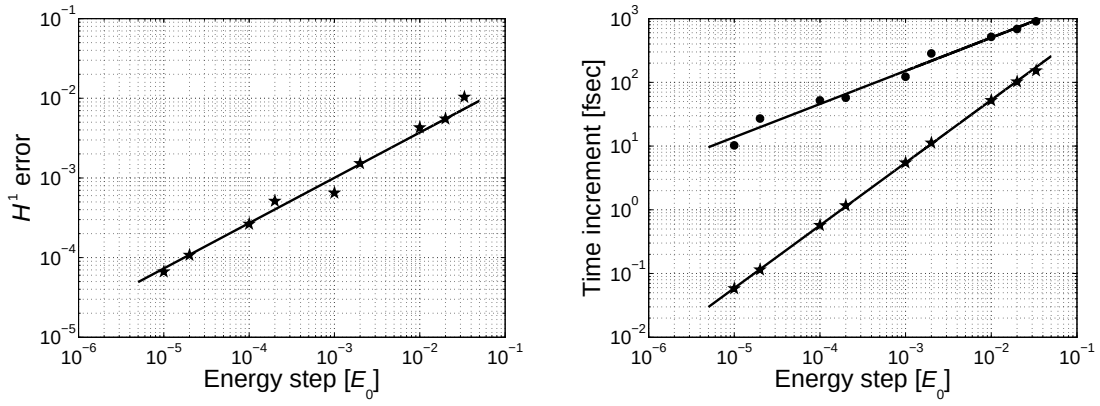


Figure 2.8: Convergence analysis of the frozen argon cluster. Left: Convergence is observed in the H^1-norm with estimated convergence rate of $r \simeq 1/2$. Right: The average ($\star$) and the maximum ($\bullet$) time steps selected by energy-stepping are $O(h)$ and $O(h^{1/2})$, respectively.

in elastodynamics). For these applications, the generalized coordinates q of the system are the coordinates of the nodes in the deformed configuration of the solid. We present two numerical examples which illustrate the performance of energy-stepping in that area of application. The first example concerns the dynamics of a spinning neo-Hookean cube, the second concerns the collision of two neo-Hookean spheres. In all applications, we assume a strain-energy density of the form

$$W(F) = \frac{\lambda_0}{2} (\log J)^2 - \mu_0 \log J + \frac{\mu_0}{2} \operatorname{tr}\left(F^T F\right) \tag{2.88}$$

which describes a neo-Hookean solid extended to the compressible range. In this expression, λ_0 and μ_0 are the Lamé constants.

In problems involving contact we additionally consider the kinematic restrictions imposed by the impenetrability constraint. We recall that the admissible configuration set $\mathcal{C}$ of a deformable body is the set of deformation mappings which are globally one-to-one. In so-called *barrier* methods, the interpenetration constraint may be accounted for by adding the indicator function $I_{\mathcal{C}}(q)$ of the admissible set $\mathcal{C}$ to the energy of the solid. We recall that the indicator function of a set $\mathcal{C}$ is the extended-valued function

$$I_{\mathcal{C}}(q) = \begin{cases} 0 & \text{if } q \in \mathcal{C} \\ \infty & \text{otherwise} \end{cases} \tag{2.89}$$

Kane *et al.* [41] and Pandolfi *et al.* [68] have provided a computationally convenient characterization of the admissible set $\mathcal{C}$ for polyhedra, such as result from a discretization of the domain $\mathcal{B}$ by simplices, as the set of configurations that are free of intersections between any pair of element faces, Figure 2.9. Thus, $q \in \mathcal{C}$ if no pair of element faces intersect, $q \notin \mathcal{C}$ otherwise. Often in calculations, the indicator function $I_{\mathcal{C}}$ is replaced by a penalty approximation $I_{\mathcal{C},\epsilon} \geq 0$ parameterized by a small parameter $\epsilon > 0$ and such that $I_{\mathcal{C},\epsilon} = 0$ over $\mathcal{C}$. In this approach, as $\epsilon \to 0$, $I_{\mathcal{C},\epsilon} \to I_{\mathcal{C}}$ pointwise and interpenetration is increasingly penalized. A convenient choice of penalty energy function for contact is of the form (Kane *et al.* [41] and Pandolfi *et al.* [68]

$$I_{\mathcal{C},\epsilon}(q) = \frac{1}{2\epsilon} \sum_{\alpha \in I} g_\alpha(q) \tag{2.90}$$

where the index set I ranges over all pairs of boundary faces and

$$g_\alpha(q) = \begin{cases} 0, & \text{if the faces do not intersect} \\ \|A - B\|^2, & \text{otherwise} \end{cases} \tag{2.91}$$

where A and B are the two extreme points of the intersection between a pair of simplices, as shown in Figure 2.9. It follows from the invariance properties of the admissible set $\mathcal{C}$ that $I_{\mathcal{C}}(q)$ and $I_{\mathcal{C},\epsilon}(q)$ are themselves invariant under the action of translations and rotations. It therefore follows that the constrained Lagrangian retains its energy and momentum preserving properties.

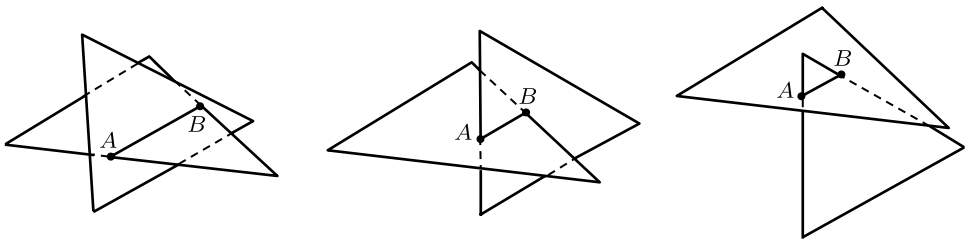


Figure 2.9: Three types of intersections between boundary simplices.

The application of energy-stepping to dynamical problems subject to set constraints, and in particular to dynamic contact problems, is straightforward. The case in which the constraints are

represented by means of a penalty energy function $I_{\mathcal{C},\epsilon}$ falls right within the general framework and requires no special considerations. In addition, energy-stepping provides an efficient means of enforcing set constraints exactly. Thus, suppose that the set constraints are accounted for by the addition of the corresponding indicator function $I_{\mathcal{C}}$ to the energy, as discussed earlier. Then, the boundary $\partial\mathcal{C}$ of the admissible set is an energy step of infinite height that necessarily causes the energy-stepping trajectory to reflect. In this case, the only modification of the algorithm that is required consists of restricting the solution of (4.39) to the admissible set $\mathcal{C}$. When the admissible set is the intersection of the zero sets of non-negative constraint functions $\{g_\alpha,\ \alpha \in I\}$, this restriction simply corresponds to choosing t_{i+1} as the minimum of the solution of (4.39) and of the solutions of

$$g_\alpha \left(q_i + (t_{i+1} - t_i)\, \dot{q}_i\right) = 0^+ \tag{2.92}$$

In particular, collision is automatically captured by the intrinsic time adaption of energy-stepping (cf., for example [10, 80], for a detailed discussion of time-step selection considerations in contact problems).

For purposes of assessing the performance of energy-stepping, in the subsequent examples we draw detailed comparisons with the second-order explicit Newmark method, namely, the member of the Newmark family of time-stepping algorithms corresponding to parameters $\beta = 0$ and $\gamma = 1/2$ (cf., for example [42] for a detailed account of Newmark's method). In the linear regime, explicit Newmark is second-order accurate and conditionally stable, with a critical time step equal to twice over the maximum natural frequency of the system. As already noted, explicit Newmark is identical to velocity Verlet. For constant time step it is also identical to central differences. It can also be shown that the Newmark solution is in one-to-one correspondence, or *shadows*, the solution of the trapezoidal-rule variational integrators (cf., for example [42]). Thus, explicit Newmark provides a convenient representative of a time-integrator commonly used in molecular dynamics, finite-differencing and variational integration.

Detailed analyses of the implicit members of the Newmark family of algorithms, their stability

and energy preserving properties (for linear systems) were given in Belytschko and Schoeberle [4], Hughes [35] and related papers.

2.5.2.1 Spinning neo-Hookean cube

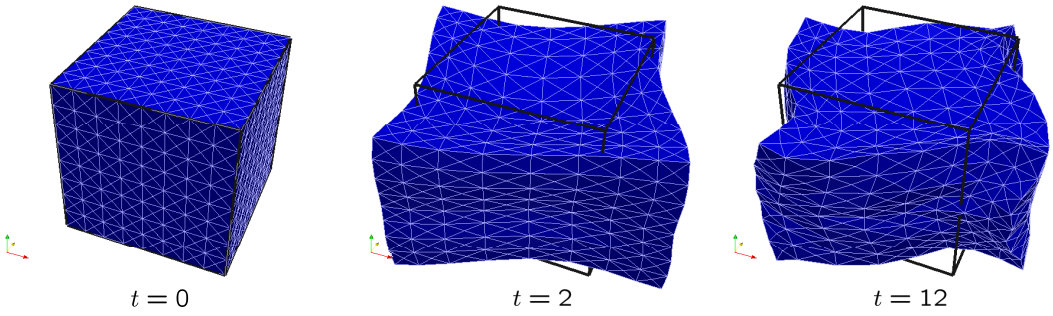


Figure 2.10: Spinning neo-Hookean cube. Snapshots of the energy-stepping trajectory for $h = 10^{-5}$ at times $t = 2$ and $t = 12$.

Our next example concerns the spinning of a free-standing elastic cube of unit size. The mesh comprises 12,288 4-node tetrahedral isoparametric elements and 2,969 nodes. The cube is a compressible neo-Hookean solid characterized by a strain-energy density of the form (2.88). The values of the material constants in (2.88) are: $\lambda_0 = 0.0100$, $\mu_0 = 0.0066$, and $\rho = 0.100$. The cube is imparted an initial angular velocity $\omega_0 = 1$ about one of its axes. The material properties are chosen such that the cube is compliant and undergoes nonlinear dynamics consisting of an overall low-frequency rotation coupled to large-amplitude high-frequency vibrations. A sequence of snapshots of the energy-stepping trajectory corresponding to $h = 10^{-5}$ are shown on the left side of Figure 2.10. The large-amplitude oscillations undergone by the spinning cube are evident in the figure.

The motion of the point $X = (0.5, 1.0, 1.0)$, the total energy, the potential energy, the kinetic energy and the z-component of the total angular momentum are shown in Figs. 2.11, 2.12, and 2.13 for the energy-stepping solutions corresponding to $h = 1 \cdot 10^{-5}$, $h = 3 \cdot 10^{-5}$ and $h = 6 \cdot 10^{-5}$, respectively, and for Newmark solutions with $\Delta t = 0.0082$, $\Delta t = 0.0204$ and $\Delta t = 0.0381$, respectively. In every case, the time step employed in the Newmark calculations is the average time step of the corresponding energy-stepping trajectory. In Figure 2.11, an ostensibly converged

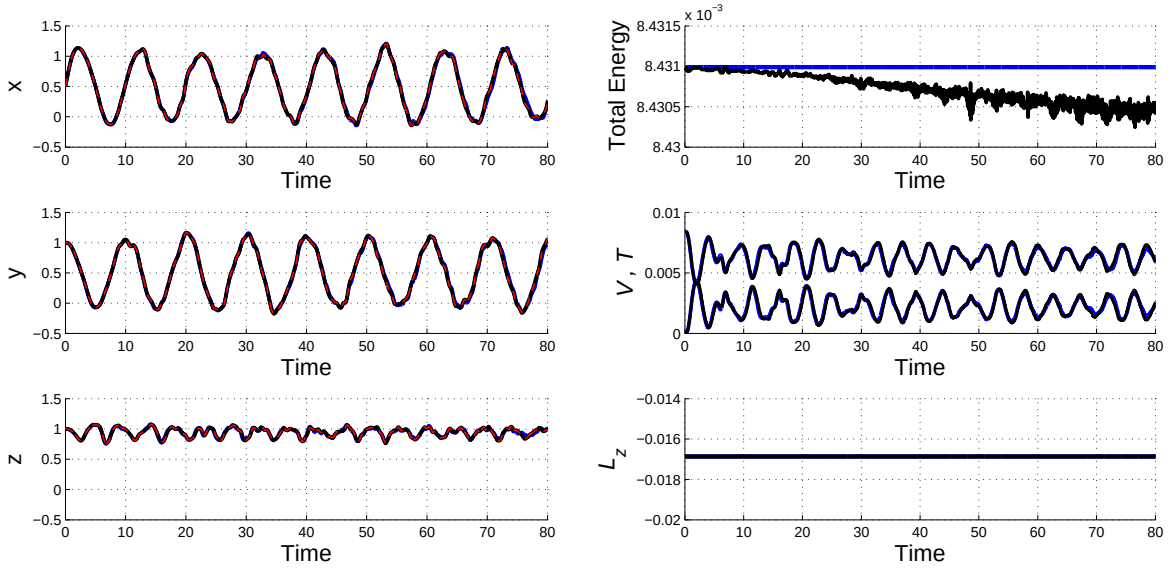


Figure 2.11: Spinning neo-Hookean cube. Blue line: Energy-stepping solution with $h = 10^{-5}$. Black line: Newmark solution with $\Delta t = 0.0082$. Red line: Newmark solution with $\Delta t = 0.0001$.

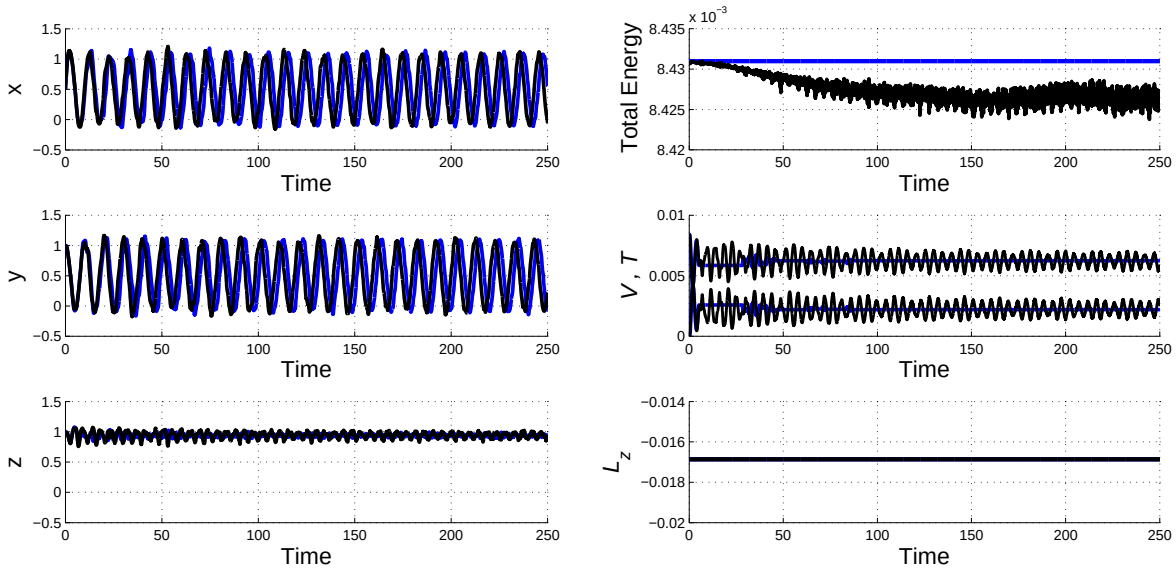


Figure 2.12: Spinning neo-Hookean cube. Blue line: Energy-stepping solution with $h = 3 \cdot 10^{-5}$. Black line: Newmark solution with $\Delta t = 0.0204$.

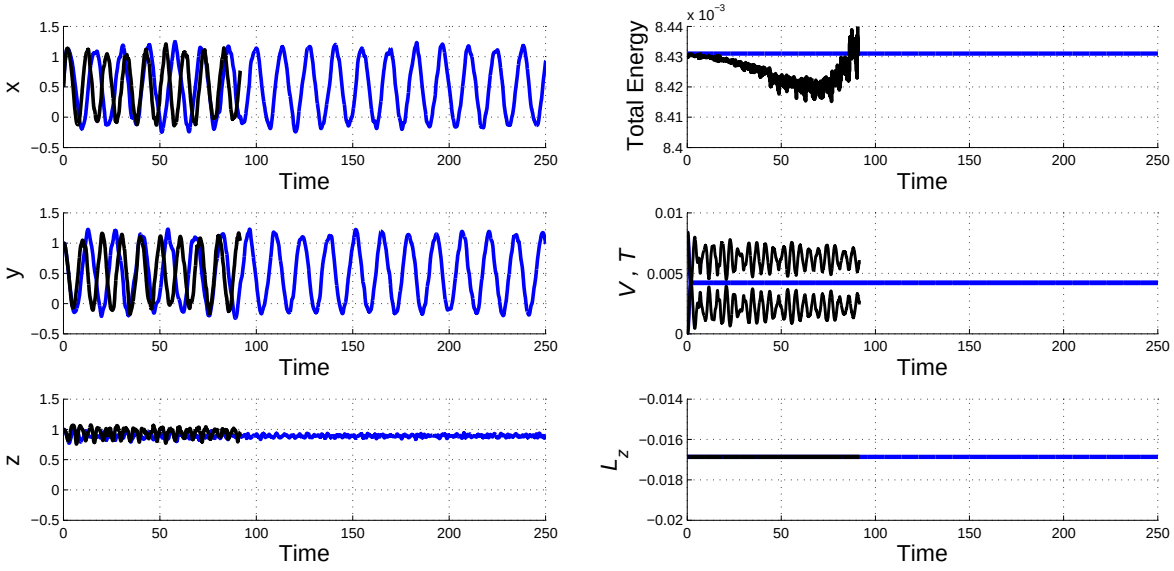


Figure 2.13: Spinning neo-Hookean cube. Blue line: Energy-stepping solution with $h = 6 \cdot 10^{-5}$. Black line: Newmark solution with $\Delta t = 0.0381$.

Newmark solution computed with a time step of $\Delta t = 0.0001$ is also shown for comparison.

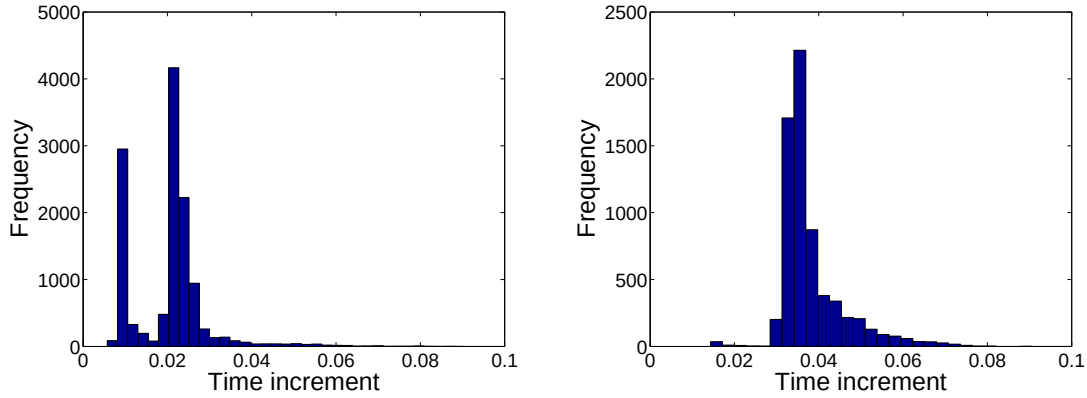


Figure 2.14: Spinning neo-Hookean cube. Histogram of time steps selected by energy-stepping. Left: $h = 3 \cdot 10^{-5}$. Right: $h = 6 \cdot 10^{-5}$.

For the smaller energy and time steps, both the energy-stepping and Newmark solutions are ostensibly converged up to a time of 80. It bears emphasis that no special precautions are taken to ensure transversality of the energy-stepping trajectory, as defined in Definition 2.4.1, which provides an indication that non-transversality is a rare event that is unlikely to turn up in practice. As expected, both solutions additionally exhibit exact linear and angular momentum conversation. In addition, the energy-stepping solution also exhibits exact energy conservation. By contrast, in

Newmark's solution the total energy, while bounded, drifts somewhat. These trends are accentuated at the intermediate energy and time steps. By way of sharp contrast, at the larger energy and time steps the stability, energy and momentum conservation properties of energy-stepping are maintained, while the Newmark solution loses stability and blows up.

Finally, Figure 2.14 shows a histogram of the time steps selected by energy-stepping for the intermediate and the largest energy steps. The broad range of time steps is noteworthy, as is the scaling of the average time step with the energy step employed in the calculation. Thus, a small (large) energy step results in comparatively smaller (larger) time steps on average, as expected. Figure 2.14 also illustrates the automatic time-selection property of energy-stepping. This property may in turn be regarded as the means by which energy-stepping achieves symplecticity and exact energy conservation. It also bears emphasis that, unlike variable time-step variational integrators designed to conserve energy [40, 51], energy-stepping always selects a valid (albeit possibly infinite) time step and is therefore free of solvability concerns.

2.5.2.2 Dynamic contact of two neo-Hookean spherical balls

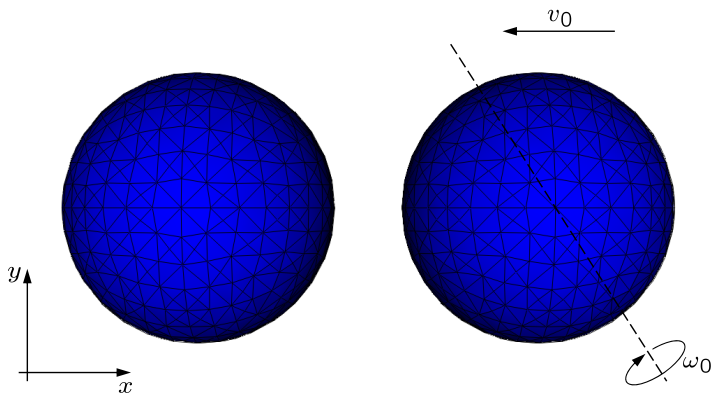


Figure 2.15: Collision of two neo-Hookean spherical balls.

Our last example concerns the collision of two free-standing elastic balls of unit radius, Figure 2.15. The mesh of each ball comprises 864 4-node tetrahedral isoparametric elements and 250 nodes. The ball is a compressible neo-Hookean solid characterized by a strain-energy density of the form (2.88). Initially, one of the balls is stationary, whereas the other ball is imparted an initial

head-on velocity of magnitude 5 and and angular velocity of magnitude $\omega_0 = 2.5$ about an axis perpendicular to the relative position vector between the centers of the balls. The values of the material constants in (2.88) are $\lambda_0 = 1.154 \cdot 10^8$, $\mu_0 = 7.96 \cdot 10^7$, and $\rho = 7,800$. As in the preceding example, the material properties are chosen such that the balls are compliant and undergo nonlinear dynamics consisting of an overall rigid-body motion coupled to large-amplitude high-frequency vibrations. The contact constraint is enforced using the penalty energy (2.90) and (2.91) with $\epsilon = 10^{-7}$.

Figure 2.16 compares the time histories of total energy, potential energy, kinetic energy, and time step attendant to the energy-stepping trajectory for $h = 10^3$ and Newmark's trajectory for $\Delta t = 3 \cdot 10^{-5}$, the latter chosen to be within the range of stability of the Newmark method. As expected, the kinetic energy of the system is partly converted to potential energy during the approach part of the collision sequence, and vice versa during the release part. We recall that period elongation, causing solutions to lag in time, is indeed a principal measure of the loss of accuracy of Newmark's method with increasing time step (cf., e. g., [35]). In this regard it is interesting to note that, despite the selection of a time step much smaller than the average energy-stepping time step, the Newmark kinetic and potential energies lag behind the corresponding energy-stepping energies. Also characteristically, energy-stepping is observed to conserve energy exactly through the collision, whereas the Newmark energy history, while bounded, drifts somewhat. Other notable features of the calculations are the ability of energy-stepping to detect the time of collision (~ 0.1) and to automatically modulate the time-step so as to resolve the fine structure of the intricate interactions that occur through the collision.

Figure 2.17 shows the time history of the center-of-mass linear and angular velocities of each of the balls and of the entire system. The center-of-mass linear and angular velocities of a system are defined as the total linear and angular momenta divided by the total mass, respectively. Because the system is free of external forces, the total linear and angular momentum of the system is conserved through the collision. We note from Figure 2.17 that the energy-stepping trajectory does indeed conserve total linear and angular momentum exactly. We also note that the transfer of linear momentum from incoming to target ball is largely in the direction of impact, with slight amounts

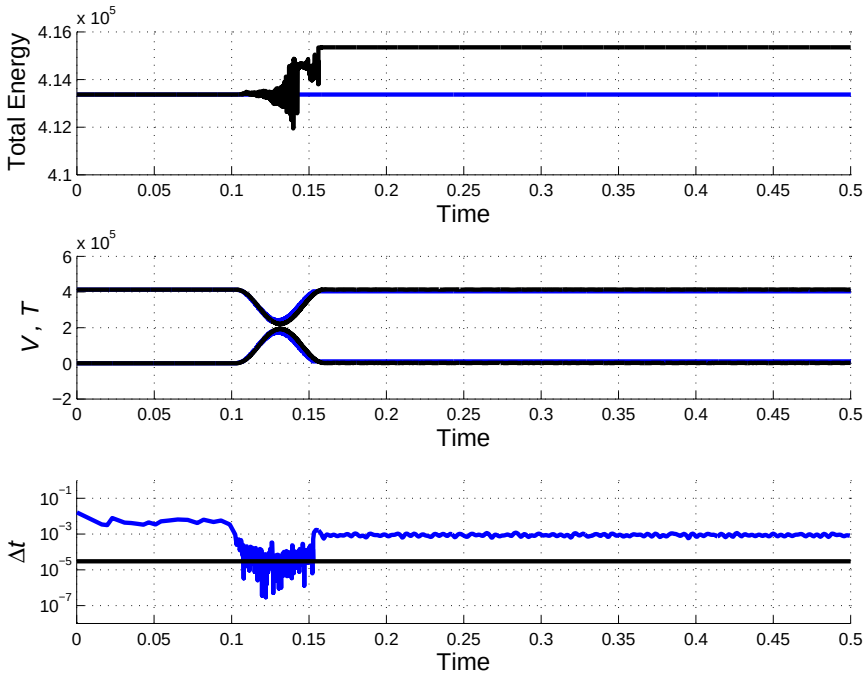


Figure 2.16: Collision of two neo-Hookean spherical balls. Blue line: Energy-stepping solution with $h = 10^3$. Black line: Newmark solution with $\Delta t = 3 \cdot 10^{-5}$.

of transfer of linear momentum in the transverse direction and of angular momentum owing to asymmetries in the mesh.

It is interesting to compare the results of the finite-element calculations with the classical theories of impact between smooth rigid bodies (see, e. g., reference [22]) and of Hertzian contact between elastic bodies. For two rigid spheres undergoing head-on impact, with the target sphere initially stationary and the incoming sphere having initial speed v_0 and angular velocity ω_0, a conventional way of expressing the final linear and angular velocities is in terms of the coefficient of restitution e. For two spheres of the same mass one has

$$\frac{T_f}{T_0} = \frac{1}{2}(1 + e^2) \tag{2.93}$$

where T_0 and T_f are the initial and final kinetic energies of the center-of-mass motion. For the

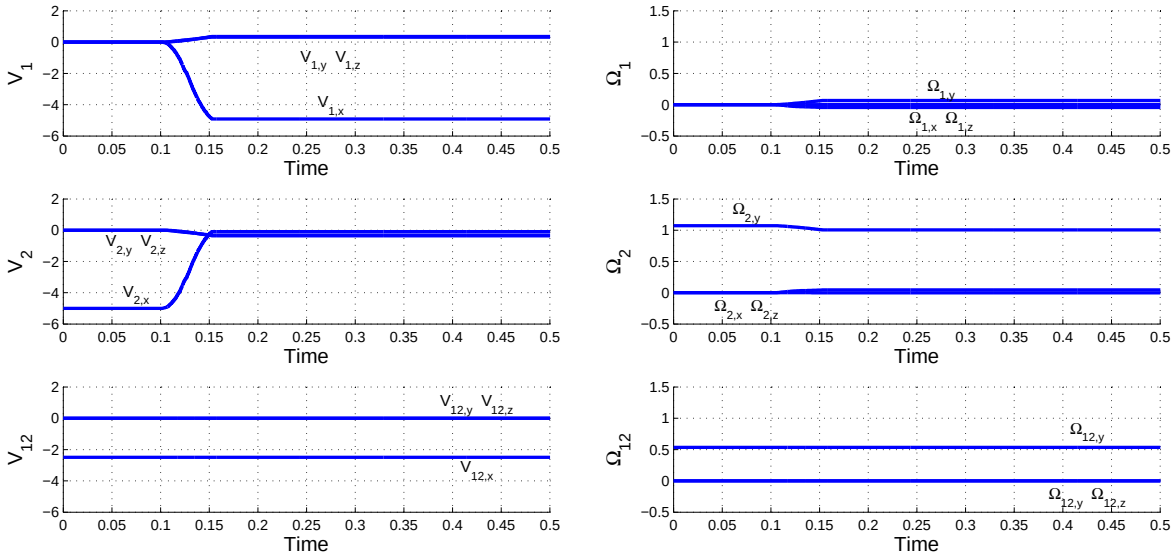


Figure 2.17: Collision of two neo-Hookean spherical balls. Energy-stepping solution with $h = 10^3$.

energy-stepping trajectory the final center-of-mass velocities of the balls are

$$V_{1,f} = (-4.91, \quad 0.31, \quad 0.35) \qquad , \qquad \Omega_{1,f} = (0.001, 0.067, -0.044) \tag{2.94a}$$

$$V_{2,f} = (-0.09, -0.31, -0.35) \qquad , \qquad \Omega_{2,f} = (-0.001, 1.005, 0.044) \tag{2.94b}$$

which roughly corresponds to a coefficient of restitution of $e = 0.964$. For elastic collisions, the coefficient of restitution provides a simple measure of the fraction of translational kinetic energy that is transferred into vibrational energy along a specific trajectory. We note, however, the coefficient of restitution is not constant but varies from trajectory to trajectory.

Another useful reference point is provided by the Hertzian theory of elastic contact (cf., e. g., [22]). While the theory applies to bodies in static equilibrium only, the resulting laws of interaction are often used to describe the dynamics of systems of spheres [22]. When applied in this manner, the theory predicts a contact duration time

$$\tau = 4.53 \left[\frac{1-\nu^2}{E} \frac{4}{3} R^3 \rho \right]^{2/5} \left[\frac{2}{v_0 R} \right]^{1/5} \tag{2.95}$$

where E is Young's modulus, ν Poisson's ratio, ρ is the mass density, R is the radius of the spheres

and v_0 is the impact velocity. For the problem under consideration, this formula gives $\tau = 0.069$, whereas the value computed from the energy-stepping solution is $\tau = 0.0476$. Thus, while Hertzian theory does provide a rough estimate of the contact time an accurate interpretation of experiments may require more detailed analyses such as presented here.

2.6 Summary and discussion

We have formulated a new class of time-integration schemes for Lagrangian mechanics, which we refer to as *energy-stepping*, that are momentum and energy conserving, symplectic and convergent. In order to achieve these properties we adopt a strategy that may be viewed as the reverse of backward-error analysis. Thus, whereas backward-error analysis seeks to identify a nearby Lagrangian system that is solved exactly by the solutions generated by a numerical integrator, the approached followed here is to directly replace the system by a nearby one that can be solved exactly. We have specifically investigated piecewise constant approximations of the potential energy obtained by replacing the original potential energy by a stepwise or *terraced* approximation at steps of uniform height. By taking steps of diminishing height, an approximating sequence of energies is generated. The trajectories of the resulting approximating Lagrangians can be characterized explicitly and consist of intervals of piecewise rectilinear motion. We have shown that the energy-stepping trajectories are symplectic, exactly conserve all the momentum maps of the original system and, subject to a transversality condition, converge to trajectories of the original system when the energy step is decreased to zero. These properties are born out by selected examples of application, including the dynamics of a frozen Argon cluster, the spinning of an elastic cube and the collision of two elastic spheres. These examples additionally showcase the excellent long-term behavior of energy-stepping, its automatic time-step selection property, and the ease with which it deals with systems with constraints, including contact problems.

It is suggestive to note that in classical Hamiltonian mechanics energy and time are conjugate variables. Conventional integration schemes are based on discretizing time, and the energy history then follows as a corollary to the integration scheme. By contract, energy-stepping may be regarded

as the result of discretizing energy, with the time increments then following from the integration scheme. In this manner, energy-stepping schemes may be regarded as *dual* to time-stepping schemes.

We close by pointing out some limitations of our analysis and possible avenues for extensions of the approach.

Firstly, our convergence analysis relies on a technical condition of transversality of the energy-stepping trajectory. It is possible that such a transversality condition can be relaxed if slight adjustments of the initial conditions are allowed and by assuming suitable lower bounds on the curvature of the energy surface. Indeed, our experience with selected numerical tests appears to indicate that the transversality condition is not of major concern in practice. However, a more clear delineation of essential and inessential conditions for the convergence of energy-stepping is desirable, if beyond the scope of this work.

Secondly, it is clear that a piecewise energy approximation of the potential energy is not the only—perhaps even the best—approximation that generates exactly solvable Lagrangians. A case in point consists of piecewise linear approximations of the potential energy over a simplicial grid. The corresponding approximating trajectories are piecewise parabolic and correspond to *free fall* within each of the simplices of the energy grid. Since the corresponding force field is piecewise constant the time-integration scheme thus defined may be thought of as *force-stepping*. The convergence properties of force-stepping are in fact much more readily established that those of energy stepping. However, a drawback of force stepping is that the approximating Lagrangian necessarily breaks some of the symmetries of the original Lagrangian in general. The systematic investigation of approximation schemes of the type proposed here, the elucidation of their properties and the determination of the best types of approximating Lagrangians in each area of application, are worthwhile directions of future research.

Appendix: Verification of the symplecticity of energy-stepping integrators

We proceed to verify that the identities (2.52) are identically satisfied by the energy-stepping scheme. For simplicity of notation and without loss of generality we take the mass matrix to be of the form $M = mI$. In addition, it suffices to consider mappings defined by the trajectory depicted in Fig. 2.2, the general result then following by recursion. Evidently, the symplectic form is trivially conserved for $t \in [t_0, t_1)$, where t_1 is the time of intersection with the potential energy jump. For $t \in (t_1, t_2)$, the dependence of t_1 on the initial conditions $(\dot{q}_0, q_0)$ must be carefully accounted for. The following relations are readily computed:

$$\begin{aligned}
\frac{\partial t_1}{\partial \dot{q}_0} &= (t_1 - t_0)\frac{\partial t_1}{\partial q_0} = -\frac{t_1 - t_0}{\dot{q}_0 \cdot n_1} n_1 = -\left|\frac{\partial t_1}{\partial \dot{q}_0}\right| n_1 \\
\frac{\partial n_1}{\partial \dot{q}_0} &= (t_1 - t_0)\frac{\partial n_1}{\partial q_0} \\
\frac{\partial n_1}{\partial q_0} &= \frac{D^2 V_1}{\|\nabla V_1\|} - \frac{n_1 \otimes n_1 \cdot D^2 V_1}{\|\nabla V_1\|} - \frac{D^2 V_1 \cdot \dot{q}_0 \otimes n_1}{\|\nabla V_1\|\,(\dot{q}_0 \cdot n_1)} + \frac{\left(n_1 \cdot D^2 V_1 \cdot \dot{q}_0\right) n_1 \otimes n_1}{\|\nabla V_1\|\,(\dot{q}_0 \cdot n_1)}
\end{aligned}$$

For $t \in (t_1, t_2)$ we have

$$\begin{aligned}
\dot{q}(q_0, \dot{q}_0, t) &= \dot{q}_0 + \lambda_1 n_1 \\
q(q_0, \dot{q}_0, t) &= q_0 + (t - t_0)\,\dot{q}_0 + (t - t_1)\,\lambda_1 n_1
\end{aligned}$$

where $\lambda_1 = \lambda_1(q_0, \dot{q}_0) = -\dot{q}_0 \cdot n_1 + \operatorname{sign}(\Delta V)\sqrt{(\dot{q}_0 \cdot n_1)^2 - \frac{2\Delta V}{m}}$ and

$$\begin{aligned}
\frac{\partial \lambda_1}{\partial \dot{q}_0} &= \frac{-\lambda_1}{\dot{q}_0 \cdot n_1 + \Delta \dot{q}} n_1 + (t_1 - t_0)\,\frac{\partial \lambda_1}{\partial q_0} \\
\frac{\partial \lambda_1}{\partial q_0} &= -\frac{\lambda_1}{\|\nabla V_1\|\,(\dot{q}_0 \cdot n_1 + \lambda_1)} \dot{q}_0 \cdot D^2 V_1 + \frac{\lambda_1\,(\dot{q}_0 \cdot n_1)}{\|\nabla V_1\|\,(\dot{q}_0 \cdot n_1 + \lambda_1)} n_1 \cdot D^2 V_1 \\
&\quad + \frac{\lambda_1\left(\dot{q}_0 \cdot D^2 V_1 \cdot \dot{q}_0\right) - \lambda_1\left(n_1 \cdot D^2 V_1 \cdot \dot{q}_0\right)(\dot{q}_0 \cdot n_1)}{(\dot{q}_0 \cdot n_1 + \lambda_1)\,\|\nabla V_1\|\,(\dot{q}_0 \cdot n_1)} n_1
\end{aligned}$$

From these identities, the components of the Jacobian matrix $T\varphi$ are found to be

$$
\begin{aligned}
P_q &= \frac{\lambda_1}{\|\nabla V_1\|} D^2 V_1 - \frac{\lambda_1 \lambda_1}{\|\nabla V_1\| (\dot{q}_0 \cdot n_1 + \lambda_1)} n_1 \otimes n_1 \cdot D^2 V_1 - \frac{\lambda_1}{\|\nabla V_1\| (\dot{q}_0 \cdot n_1 + \lambda_1)} n_1 \otimes \dot{q}_0 \cdot D^2 V_1 \\
&\quad - \frac{\lambda_1}{\|\nabla V_1\| (\dot{q}_0 \cdot n_1)} D^2 V_1 \cdot \dot{q}_0 \otimes n_1 + \frac{\lambda_1 (\dot{q}_0 + \lambda_1 n_1) \cdot \left(D^2 V_1 \cdot \dot{q}_0\right)}{(\dot{q}_0 \cdot n_1 + \lambda_1) \|\nabla V_1\| (\dot{q}_0 \cdot n_1)} n_1 \otimes n_1 \\
Q_q &= (t - t_1) P_q + \left[I + \frac{\lambda_1}{\dot{q}_0 \cdot n_1} n_1 \otimes n_1 \right] \\
P_p &= I - \frac{\lambda_1}{(\dot{q}_0 \cdot n_1 + \lambda_1)} n_1 \otimes n_1 + (t_1 - t_0) P_q \\
Q_p &= (t - t_1) P_p + (t_1 - t_0) \left[I + \frac{\lambda_1}{\dot{q}_0 \cdot n_1} n_1 \otimes n_1 \right]
\end{aligned}
$$

In addition,

$$
\begin{aligned}
P_p^T Q_p &= (t - t_1) P_p^T P_p + (t_1 - t_0) I + (t_1 - t_0)^2 P_q^T \cdot \left[I + \frac{\lambda_1}{\dot{q}_0 \cdot n_1} n_1 \otimes n_1 \right] = sym \\
Q_q^T P_q &= (t - t_1) P_q^T P_q + \left[I + \frac{\lambda_1}{\dot{q}_0 \cdot n_1} n_1 \otimes n_1 \right] \cdot P_q = sym \\
P_p^T Q_q &= Q_p^T P_q + I = I + (t - t_1) \left[I - \frac{\lambda_1}{\dot{q}_0 \cdot n_1 + \lambda_1} n_1 \otimes n_1 \right] \cdot P_q \\
&\quad + (t_1 - t_0)(t - t_1) P_q^T P_q + (t_1 - t_0) P_q^T \cdot \left[I + \frac{\lambda_1}{\dot{q}_0 \cdot n_1} n_1 \otimes n_1 \right]
\end{aligned}
$$

and the symplecticity relations (2.52) follow from the identity

$$
\begin{aligned}
\left[I + \frac{\lambda_1}{\dot{q}_0 \cdot n_1} n_1 \otimes n_1 \right] \cdot P_q = sym &= \frac{\lambda_1}{\|\nabla V_1\|} D^2 V_1 + \frac{\lambda_1 \left(\dot{q}_0 \cdot D^2 V_1 \cdot \dot{q}_0\right)}{\|\nabla V_1\| (\dot{q}_0 \cdot n_1)^2} n_1 \otimes n_1 \\
&\quad - \frac{\lambda_1}{\|\nabla V_1\| (\dot{q}_0 \cdot n_1)} \left[D^2 V_1 \cdot \dot{q}_0 \otimes n_1 + n_1 \otimes \dot{q}_0 \cdot D^2 V_1 \right]
\end{aligned}
$$

The symplecticity of the energy-stepping scheme is thus directly verified.

Chapter 3

Force-stepping integrators

We formulate an integration scheme for Lagrangian mechanics, referred to as the *force-stepping scheme*, that is symplectic, energy conserving, time-reversible and convergent with automatic selection of the time step size. The scheme also conserves approximately all the momentum maps associated with the symmetries of the system. Exact conservation of momentum maps may additionally be achieved by recourse to Lagrangian reduction. The force-stepping scheme is obtained by replacing the potential energy by a piecewise affine approximation over a simplicial grid, or regular triangulation. By taking triangulations of diminishing size, an approximating sequence of energies is generated. The trajectories of the resulting approximate Lagrangians can be characterized explicitly and consist of piecewise parabolic motion, or *free fall*. Selected numerical tests demonstrate the excellent long-term behavior of force-stepping, its automatic time-step selection property, and the ease with which it deals with constraints, including contact problems.

3.1 Introduction

In a recent paper [24] and in Chapter 2, the authors have proposed a method of approximation for Lagrangian mechanics consisting of replacing the Lagrangian $L(q, \dot{q})$ of the system by a sequence of approximate Lagrangians $L_h(q, \dot{q})$ that can be *solved exactly*. The approximate solutions $q_h(t)$ are then the *exact* trajectories of $L_h(q, \dot{q})$. In this manner, the approximate solutions are themselves trajectories of a Lagrangian system and, therefore, have many of the properties of such trajectories such as symplecticity and exact energy conservation. If, in addition, the approximation of the Lagrangian preserves its symmetry group, then the approximate trajectories conserve all the momentum maps of the system. It bears emphasis that, in contrast to variational integrators (cf., e. g., [52, 53, 58]), the momenta that are conserved are the exact momenta of the system and not some time discretization thereof. Finally, if the approximate Lagrangians $L_h(q, \dot{q})$ converge to the exact one $L(q, \dot{q})$ in a topology that implies, in particular, convergence of stationary points, then the approximate trajectories $q_h(t)$ converge to exact trajectories $q(t)$ of the original system.

In [24], this program has been carried out in full for Lagrangians of the type: $L(q, \dot{q}) = K(\dot{q}) - V(q)$, with $K(\dot{q})$ a quadratic form, and for approximate Lagrangians of the type $L_h(q, \dot{q}) = K(\dot{q}) - V_h(q)$, with $V_h(q)$ a terraced piecewise constant approximation to $V(q)$, Figure 2.1b. This type of approximation does indeed result in exactly solvable Lagrangians. The approximate trajectories are piecewise rectilinear, with the intervals of rectilinear motion spanning consecutive level contours of the potential. Conveniently, the terraced potential $V_h(q)$ has all the symmetries of the original potential $V(q)$, and the approximate trajectories exactly conserve all the momentum maps of the system, whether explicitly known or not. The durations of the intervals of rectilinear motion may be regarded as time steps, whose determination is part of the solution process. In this manner, the approach overcomes an intrinsic limitation of fixed time-step variational integrators, which cannot simultaneously conserve energy, the symplectic structure and other conserved quantities, such as linear and angular momenta [19]. Under mild restrictions on the potential, the approximate trajectories are found to indeed converge to exact trajectories of the system, subject to technical transversality constraints.

The terraced piecewise constant approximation of the potential considered in [24] may be thought of as the lowest-order approximation that results in convergence. Evidently, approximations of increasing order can be obtained by recourse to piecewise polynomial interpolation of the potential. In this chapter, we consider the next order of approximation consisting of approximate potentials $V_h(q)$ obtained by piecewise-linear interpolation of $V(q)$ over structured simplicial meshes, Figure 2.1c. Evidently, within each simplicial element in the interpolation mesh, the forces are constant, hence the term *force-stepping*, and the system is in *free fall*. In particular, the approximate Lagrangian can be solved exactly and the approximate trajectories are piecewise quadratic. In Section 3.3, we provide an efficient implementation of the scheme that establishes its feasibility in practical, possibly high-dimensional, applications. In particular, we show in Section 3.6 that the typical numbers of time steps, or time intervals between crossings of simplicial boundaries, and their sizes are within the expected range for explicit integration. In Section 3.5, we show that the approximate trajectories thus computed do indeed converge to exact trajectories of the system.

However, the matter of conservation requires careful attention as the piecewise-linear approximate Lagrangians break the symmetries of the system in general, which in turn results in a lack of exact momentum conservation of the force-stepping scheme. This symmetry breaking raises an interesting challenge that is not present in the case of the energy-stepping scheme and that we address in several ways. The first way to deal with the lack of exact conservation of the force-stepping scheme—and possibly the most effective—is to do nothing. Indeed, we show in Section 3.4 that the force-stepping scheme is nearly-conserving for all symmetries of the system, whether explicitly known or not. The conservation error is controlled uniformly on compact time intervals by the asymmetry of the approximate Lagrangian. The numerical experiments presented in Section 3.6 further show that the corresponding momentum maps tend to remain nearly constant uniformly for all times. This near-conservation property is frequently sufficient in applications that require good long-time behavior of the solutions. An alternative strategy, described in Section 3.4.2, for avoiding broken symmetries in the force-stepping scheme is to appeal to the theory of Lagrangian reduction. For symmetries for which explicit reduced Lagrangians are known, an application of the force-stepping scheme to

the reduced Lagrangian results in exact momentum conservation. A case in point is translational invariance, which can be dealt with effectively by the introduction of Jacobi and center of mass coordinates.

The chapter is organized as follows. The force-stepping time-integration scheme is defined in Section 3.2 and the construction of continuous piecewise-linear approximate energies is discussed in Section 3.3. The conservation properties of force-stepping, together with a general procedure for reducing translational symmetry, are presented in Section 3.4. In Section 3.5 we prove convergence of trajectories except for a negligibly small set of initial conditions. In Section 3.6 we present selected examples of application that illustrate the convergence, accuracy and conservation properties of force-stepping, including the motion of two bodies which attract each other by Newtons law of gravitation, that is the Kepler problem, the dynamics of a frozen argon cluster, and the oblique impact of an elastic cube. Finally, a summary and concluding remarks are collected in Section 3.7.

3.2 Force-stepping integrators

For definiteness, we specifically consider dynamical systems characterized by Lagrangians $L : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ of the form

$$L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V(q) \tag{3.1}$$

where M is the mass matrix and V is the potential energy. Lagrangians of this form arise in a number of areas of application including structural dynamics, celestial mechanics and molecular dynamics. The trajectories of a Lagrangian system can be approximated by replacing $L(q, \dot{q})$ by an approximating Lagrangian $L_h(q, \dot{q})$ that can be solved exactly. A particular type of approximating Lagrangian is

$$L_h(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V_h(q) \tag{3.2}$$

obtained by introducing an approximation of the potential energy. In this work, we specifically investigate continuous piecewise-linear approximations of the potential energy. Thus, we construct a regular triangulation $\mathcal{T}_h$ of $\mathbb{R}^d$, as shown in Section 3.3, and define V_h as the corresponding

continuous piecewise affine approximation of V. By taking triangulations of diminishing size, an approximating sequence of energies and Lagrangians is generated in this manner.

The chief characteristics of the new systems thus obtained are that they can be solved exactly, as demonstrated in Section 3.2.1, and that they retain symmetries of the original system, as shown in Section 3.4. In contrast to piecewise constant interpolations of the potential energy, which preserve all the symmetries of the system [24], piecewise-linear interpolations break some symmetries in general. This difficulty can be partially overcome by recourse to Lagrangian reduction, as demonstrated in Section 3.4. However, we emphasize that the only assumption on the Lagrangian is its form—Equation (3.1)—and therefore no further considerations to Lagrangian reduction will be required in this section, i.e., $L(q, \dot{q})$ may also be a reduced Lagrangian.

3.2.1 Computation of the exact trajectories of the approximating Lagrangian

We now proceed to explicitly compute the exact trajectories of the approximate Lagrangian $L_h(q, \dot{q})$ resulting from V_h, the continuous piecewise-linear approximation of the potential energy. Suppose that the system is in configuration q_0 at time t_0 and in configuration q_2 at time t_2 and that during the time interval $[t_0, t_2]$ the system intersects one single boundary B_1 separating two regions Δ_I and Δ_{II}, of the underlying triangulation, with linear energies $V_I + \nabla V_I \cdot (q - q_I)$ and $V_{II} + \nabla V_{II} \cdot (q - q_{II})$, Figure 3.1. By the construction of V_h, the following continuity condition is attained at the boundary surface B_1

$$V_I + \nabla V_I \cdot (q - q_I) = V_{II} + \nabla V_{II} \cdot (q - q_{II}) \qquad \forall q \in B_1 \tag{3.3}$$

For simplicity, we shall further assume that V is differentiable and that all boundary crossings are transversal, i.e.,

$$n(q_1) \cdot \dot{q}_1 \neq 0 \tag{3.4}$$

where $n(q_1)$ is a vector normal to B_1 pointing in the direction of advance. It is possible for discrete trajectories to be non-transversal and therefore become ambiguously defined. However, the set of

initial conditions which result in non-transversal trajectories has negligible dimension and will not be given further consideration. More precisely, we will show in Section 3.5 that this exceptional set has Hausdorff-dimension $2d-1$ in the $2d$-dimensional phase space and consequently its Lebesgue measure vanishes.

Under the preceding assumptions, the action integral over the time interval $[t_0, t_2]$ follows as

$$I_h = \int_{t_0}^{t_2} L_h\left(q, \dot{q}\right)\ dt = \int_{t_0}^{t_1} L_h(q, \dot{q})\ dt + \int_{t_1}^{t_2} L_h(q, \dot{q})\ dt \tag{3.5}$$

where t_1 is the time at which the trajectory intersects B_1. In regions where $V_h(q)$ is linear the trajectory $q(t)$ is quadratic in time. Therefore, the action of the system can be computed exactly and reduces to

$$\begin{aligned} I_h = (t_1 - t_0)\left\{\frac{1}{2}\left(\frac{q_1 - q_0}{t_1 - t_0}\right)^T M \left(\frac{q_1 - q_0}{t_1 - t_0}\right)\right\} + (t_2 - t_1)\left\{\frac{1}{2}\left(\frac{q_2 - q_1}{t_2 - t_1}\right)^T M \left(\frac{q_2 - q_1}{t_2 - t_1}\right)\right\} - \\ (t_1 - t_0)\left\{V_I + \frac{1}{2}\left(q_1 + q_0 - 2q_I\right)\cdot\nabla V_I + \frac{(t_1 - t_0)^2}{24}\nabla V_I^T M^{-1}\nabla V_I\right\} - \\ (t_2 - t_1)\left\{V_{II} + \frac{1}{2}\left(q_1 + q_2 - 2q_{II}\right)\cdot\nabla V_{II} + \frac{(t_2 - t_1)^2}{24}\nabla V_{II}^T M^{-1}\nabla V_{II}\right\} \end{aligned} \tag{3.6}$$

where $q_1 = q\left(t_1\right)$ is constrained to be on the boundary surface B_1. Stationarity of the action with respect to (t_1, q_1) additionally gives the energy conservation equation

$$\begin{aligned} \left(\frac{q_1 - q_0}{t_1 - t_0}\right)^T M \left(\frac{q_1 - q_0}{t_1 - t_0}\right) + (q_0 - q_1)\cdot\nabla V_I + \frac{(t_1 - t_0)^2}{4}\nabla V_I^T M^{-1}\nabla V_I = \\ \left(\frac{q_2 - q_1}{t_2 - t_1}\right)^T M \left(\frac{q_2 - q_1}{t_2 - t_1}\right) + (q_2 - q_1)\cdot\nabla V_{II} + \frac{(t_2 - t_1)^2}{4}\nabla V_{II}^T M^{-1}\nabla V_{II} \end{aligned} \tag{3.7}$$

and the linear momentum balance equation

$$M\frac{q_1 - q_0}{t_1 - t_0} - \frac{1}{2}\left(t_1 - t_0\right)\nabla V_I - M\frac{q_2 - q_1}{t_2 - t_1} - \frac{1}{2}\left(t_2 - t_1\right)\nabla V_{II} + \lambda n(q_1) = 0 \tag{3.8}$$

where λ is a Lagrange multiplier.

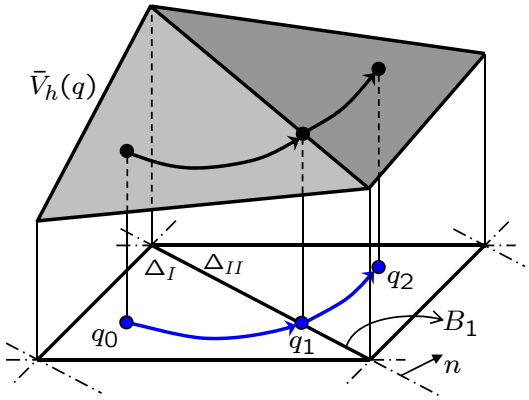


Figure 3.1: Trajectory of a system whose potential energy is approximated as continuous piecewise-linear.

In order to make a more direct connection with time-integration schemes we reformulate the problem slightly by assuming that t_0, q_0—the latter on a boundary surface B_0 except, possibly, at the initial time—and the initial velocity

$$\dot{q}_0^+ = \dot{q}\left(t_0^+\right) = \frac{q_1 - q_0}{t_1 - t_0} + \frac{1}{2}\left(t_1 - t_0\right) M^{-1} \nabla V_I \tag{3.9}$$

are known. Let t_1 and q_1 be the time and point at which the trajectory intersects the next boundary surface B_1. We then seek to determine

$$\dot{q}_1^+ = \dot{q}\left(t_1^+\right) = \frac{q_2 - q_1}{t_2 - t_1} + \frac{1}{2}\left(t_2 - t_1\right) M^{-1} \nabla V_{II} \tag{3.10}$$

A reformulation of eqs. (3.7) and (3.8) in terms of $\dot{q}_1^+$ gives

$$\left(\dot{q}_1^+\right)^T M \dot{q}_1^+ = \left(\dot{q}_0^+\right)^T M \dot{q}_0^+ - 2\left(q_1 - q_0\right) \nabla V_I \tag{3.11}$$

$$\dot{q}_1^+ = \dot{q}_0^+ - \left(t_1 - t_0\right) \nabla V_I + \lambda M^{-1} n(q_1) \tag{3.12}$$

Then, the system (3.11 - 3.12) has the solution

$$\dot{q}_1 = \dot{q}\left(t_1^+\right) = \dot{q}\left(t_1^-\right) = \dot{q}_0^+ - (t_1 - t_0)\, M^{-1}\nabla V_I \tag{3.13}$$

3.2.2 Stationary points of the action integral as weak solutions of the equation of motion

According to Hamilton's principle, the physical trajectories of a smooth Lagrangian system are the stationary points of the action integral $I = \int_0^T L_h(q_h(t), \dot{q}_h(t))\, dt$, $T > 0$. By analogy, the trajectories of an approximating Lagrangian L_h of the form (3.2) should emerge as critical points of the corresponding action functional I_h. We note that in the case of a continuous piecewise-linear potential energy the Lagrangian L_h is not differentiable everywhere and so the notion of critical points does not have a well-defined meaning in the classical sense. However, the calculations in the previous subsection can be justified by interpreting stationary points of the action integral I_h as weak solutions of the equation of motion. As we will show below in Section 3.5, for almost all initial conditions a discrete trajectory will not spend a positive amount of time in $\partial\mathcal{T}_h$, the union of element boundaries in $\mathcal{T}_h$. In order to derive the equation of motion, we will therefore restrict to such trajectories.

Theorem 3.2.1 *Suppose $q_h \in W^{1,2}((0,T);\mathbb{R}^d)$ satisfies $|\{t : q_h(t) \in \partial\mathcal{T}_h\}| = 0$. Then, the action functional I_h is differentiable at q_h and q_h is a critical point of the action functional if and only if*

$$M\ddot{q}_h(t) = -\nabla V_h(q_h(t)) \tag{3.14}$$

in the sense of distributions.

Note that the right hand of (3.14) is an element of $L^\infty((0,T);\mathbb{R}^d)$, so q_h is a distributional solution of (3.14) if and only if $q_h \in W^{2,\infty}((0,T);\mathbb{R}^d)$ and $M\ddot{q}_h(t) = -\nabla V_h(q_h(t))$ holds almost everywhere.

Proof. Let $q_h, \varphi \in W^{1,2}((0,T);\mathbb{R}^d)$. Clearly, the first term in the first variation of I_h is

$$\lim_{\varepsilon\to 0}\frac{1}{2\varepsilon}\left(\int_0^T (\dot{q}_h(t)+\varepsilon\dot{\varphi}(t))^T M(\dot{q}_h(t)+\varepsilon\dot{\varphi}(t)) - \dot{q}_h^T(t)M\dot{q}_h(t)\,dt\right) = \int_0^T \dot{q}_h^T(t)M\dot{\varphi}(t)\,dt \tag{3.15}$$

On the other hand, if $|\{t : q_h(t) \in \partial\mathcal{T}_h\}| = 0$, then

$$\frac{1}{\varepsilon}\left(V_h(q_h(t)+\varepsilon\varphi(t)) - V_h(q_h(t))\right) \to \nabla V_h(q_h(t))\cdot\varphi(t) \tag{3.16}$$

almost everywhere as $\varepsilon \to 0$. Since V_h is locally Lipschitz, we furthermore have

$$\left|\frac{1}{\varepsilon}V_h(q_h(t)+\varepsilon\varphi(t)) - V_h(q_h(t))\right| \le C\|\varphi\|_{L^\infty} \le C\|\varphi\|_{W^{1,2}} \tag{3.17}$$

bounded independently of t. By dominated convergence we may therefore conclude that the second term in the first variation of I_h is

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^T \left(V_h(q_h(t)+\varepsilon\varphi(t)) - V_h(q_h(t))\right)\,dt = \int_0^T \nabla V_h(q_h(t))\cdot\varphi(t)\,dt \tag{3.18}$$

Summarizing, we have shown that the action functional is differentiable at q_h and that its derivative in direction φ is given by

$$\delta I_h(q,\varphi) = \int_0^T \dot{q}_h^T(t)M\dot{\varphi}(t) - \nabla V_h(q_h(t))\cdot\varphi(t)\,dt \tag{3.19}$$

This expression vanishes if and only if q_h is a weak solution to (3.14). ■

3.2.3 Summary of the force-stepping scheme

We close this section by summarizing the relations obtained in the foregoing and defining the *force-stepping* approximation scheme resulting from a continuous piecewise-linear approximation of the potential energy.

Definition 3.2.1 (Force-stepping) Suppose $(t_k, q_k, \dot{q}_k)$ and a continuous piecewise-linear approximation of the potential energy V_h are given. Let t_{k+1} and q_{k+1} be the time and point of boundary crossing of the parabolic trajectory $q_k + (t - t_k)\,\dot{q}_k - \frac{1}{2}\,(t - t_k)^2\,M^{-1}\nabla V_k$. Then, the updated velocity is $\dot{q}_{k+1} = \dot{q}_k - (t_{k+1} - t_k)\,M^{-1}\nabla V_k$.

These relations define a discrete propagator

$$\Phi_h : (t_k, q_k, \dot{q}_k) \mapsto (t_{k+1}, q_{k+1}, \dot{q}_{k+1}) \tag{3.20}$$

that can be iterated to generate a discrete trajectory.

Algorithm 3 Force-stepping integrator

Require: $V(q)$, q_0, $\dot{q}_0$, t_0, t_f, $\Delta_0 \in \mathcal{T}_h$, V_0 and ∇V_0
1: $k \leftarrow 0$
2: **while** $t_k < t_f$ **do**
3: $\{t_{k+1}, B_s\} \leftarrow$TIME-STEP$(q_k, \dot{q}_k, \nabla V_k, t_k; \Delta_k)$
4: $q_{k+1} \leftarrow q_k + (t_{k+1} - t_k)\,\dot{q}_k - \frac{1}{2}\,(t_{k+1} - t_k)^2\,M^{-1}\nabla V_k$
5: $\dot{q}_{k+1} \leftarrow \dot{q}_k - (t_{k+1} - t_k)\,M^{-1}\nabla V_k$
6: $V_{k+1} \leftarrow V_k + (q_{k+1} - q_k) \cdot \nabla V_k$
7: $\{\nabla V_{k+1}, \Delta_{k+1}\} \leftarrow$UPDATE$(\nabla V_k, \Delta_k; B_s)$
8: $k \leftarrow k + 1$
9: **end while**

The implementation of the force-stepping integrator is summarized in Algorithm 3. The algorithm consists of two methods. The first method TIME-STEP determines the time of exit of the parabolic trajectory from the simplex $\Delta_k \in \mathcal{T}_h$

$$q_k + (t_{k+1} - t_k)\,\dot{q}_k - \frac{1}{2}\,(t_{k+1} - t_k)^2\,M^{-1}\nabla V_k \in B_s \tag{3.21}$$

where $B_s \in \partial\Delta_k$ is the boundary surface intersected. The second method UPDATE is responsible for updating all simplex-related information, e. g. ∇V_{k+1} of the adjacent simplex Δ_{k+1}. Both tasks can be effectively accomplished upon a unique, systematic and efficient representation of the continuous piecewise-linear approximate potential energy, as presented in Section 3.3.

It bears emphasis that force-stepping requires the solution of no system of equations and, therefore, its complexity is comparable to that of explicit methods. However, the need to compute the root of a nonlinear function and to update all simplex-related information adds to the overhead of one application of the algorithm. It is still possible, however, that such overhead may be offset by the higher accuracy in particular applications. These and other trade-offs are investigated subsequently by way of numerical testing.

3.3 Continuous piecewise linear representation of the approximate potential V_h

One of the most successful applications of continuous piecewise-linear functions has been in nonlinear circuit theory, in particular, in the field of nonlinear resistive networks. In 1965, Katzenelson [45] presented an effective approach for the search of the operating point of a piecewise-linear resistive network. Since then, the same idea has been extended, improved and generalized [11,13,18]. Among these extensions and generalizations are the resolution of general nonlinear equations of the form $f(x) = 0$ where $f : \mathbb{R}^d \to \mathbb{R}^m$ is a continuous mapping [12], and the introduction of a canonical piecewise-linear function [43, 66] which allows for the compact and closed representation of any function using the minimal number of parameters by taking advantage of a simplicial partition of the domain [39].

Our aim in this section is to describe a unique, systematic and efficient representation of the continuous piecewise-linear approximation of a function $f : \mathbb{R}^d \to \mathbb{R}$, in general, and the potential energy $V : \mathbb{R}^d \to \mathbb{R}$, in particular. With this goal in mind, we first consider a region $\mathcal{S} \in \mathbb{R}^d$ and its regular tessellation in hyperrectangles or orthotopes with characteristic length $h_j \in \mathbb{R}^+$ in the j-th direction, with $j = 1, 2, \ldots, d$. We then subdivide each hyperrectangle into proper simplices by introducing a unique, regular triangulation $\mathcal{T}_h$ in accordance with Chien and Kuh [11], see Figure 3.2. For later reference, we define the homeomorphism $\Lambda : \mathcal{S} \to \mathcal{S}_C$, represented by a diagonal matrix T, i.e., $z = Tq$, which maps hyperrectangles in $\mathcal{S}$ into hypercubes $[0, 1]^d$ in a new region $\mathcal{S}_C \in \mathbb{R}^d$. For

later reference, we name $\mathcal{V}$ and $\mathcal{V}_C$ the sets of vertices contained in regions $\mathcal{S}$ and $\mathcal{S}_C$, respectively.

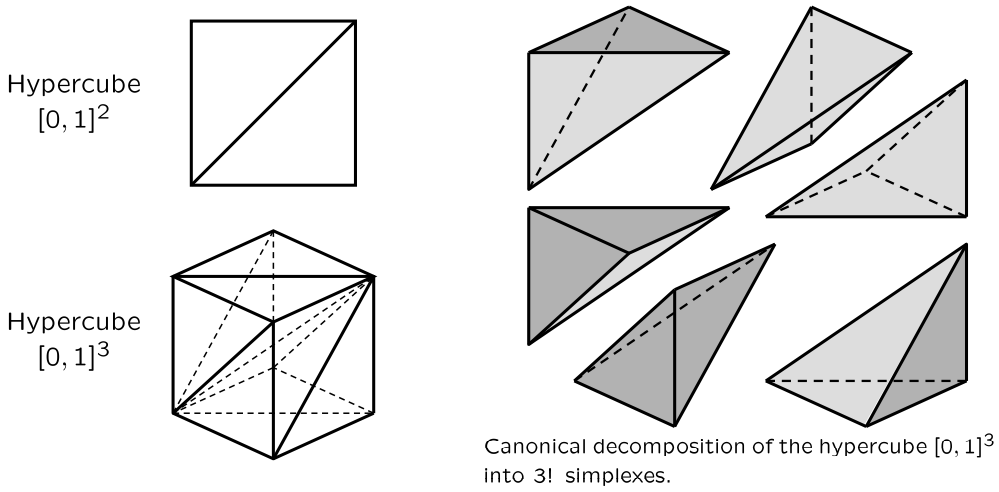


Figure 3.2: Simplicial partition of hypercubes $[0,1]^2$ and $[0,1]^3$.

Before we discuss further details of the continuous piecewise-linear representation of the approximate potential energy, we may first recall some properties of a simplex and its boundaries.

Definition 3.3.1 *Let $q^0, q^1, \ldots, q^d$ be $(d+1)$ points in general position in the d-dimensional space. A simplex $\Delta(q^0, q^1, \ldots, q^d)$ is defined as the convex hull of $q^0, q^1, \ldots, q^d$—the vertices of the simplex—*

$$\Delta(q^0, q^1, \ldots, q^d) = \left\{ q : q = \sum_{i=0}^{d} \mu_i q^i, 1 \geq \mu_i \geq 0, i = 0, 1, 2, \ldots, d \text{ and } \sum_{i=0}^{d} \mu_i = 1 \right\}$$

In addition, corresponding to the $(d+1)$ vertices, there are $(d+1)$ boundaries. The boundary B_s which corresponds to the vertex q^s is defined as

$$B_s = \left\{ q : q \in \Delta(q^0, q^1, \ldots, q^d) \text{ with } \mu_s = 0 \right\}$$

Since B_s contains all the vertices except q^s, there is a one-to-one correspondence between vertices and boundaries.

We now present the key steps for achieving a unique and systematic representation of force-stepping trajectories. To this end, we first subdivide each hypercube which belongs to $\mathcal{S}_C$ into

non-overlapping simplices by properly arranging vertices of $\mathcal{V}_C$ in a fixed order [11]. This ordering relation allows for a unique representation of every point $z \in \mathcal{S}_C$ and $q \in \mathcal{S}$ (see Lemma 3.3.1). Therefore, there is a unique representation of the affine function $V_h(q) : \mathbb{R}^d \to \mathbb{R}$ with $V_h(q^i) = V(q^i)$ for all vertices $q^i \in \mathcal{V}$ (see Lemma 3.3.2) upon which force-stepping trajectories are systematically built.

Lemma 3.3.1 *Every $z \in \mathcal{S}_C$ has a unique representation $z = \sum_{i=0}^{m} \mu_i z^i$ where $\mu_i > 0$, $z^i \in \mathcal{V}_C$, for $i = 0, 1, \ldots, m(\leq d)$, $\sum_{i=0}^{m} \mu_i = 1$, and the following ordering relation is attained $z^0 \leqslant z^1 \leqslant \ldots \leqslant z^m \leqq z^0 + 1$. Likewise, every $q \in \mathcal{S}$ has a unique representation $q = \sum_{i=0}^{m} \mu_i q^i$ where $q^i = T^{-1} z^i \in \mathcal{V}$, and the following ordering relation is attained $Tq^0 \leqslant Tq^1 \leqslant \ldots \leqslant Tq^m \leqq Tq^0 + 1$.*

Proof. The proof in $\mathcal{S}_C$ is given by Kuhn [47]. The proof in $\mathcal{S}$ follows from the fact that the homeomorphism $\Lambda(\cdot)$ is defined by a positive-definite diagonal matrix T and therefore preserves the ordering of vertices. ■

Corollary 3.3.1 *If $m = d$, z is an interior point of the simplex $\Delta(z^0, z^1, \ldots, z^d)$. Otherwise, z lies on the boundary of a simplex. In addition, every d-dimensional simplex defined by $(d+1)$ vertices contains z^0 and $z^0 + 1$ which define the hypercube $C(z^0)$ containing the simplex.*

Lemma 3.3.2 *A nonlinear function $V(q) : \mathbb{R}^d \to \mathbb{R}$ is approximated by an affine function $V_h(q) : \mathbb{R}^d \to \mathbb{R}$ as follows*

$$V_h(q) = \nabla V_\Delta \cdot q + V_0$$

for $q \in \Delta(q^0, q^1, \ldots, q^d)$, $\nabla V_\Delta = \bar{J}_{h,[1\ldots d]} T \in \mathbb{R}^d$—the first d elements of the vector—and $V_0 = \bar{J}_{h,[d+1]} \in \mathbb{R}$. The Jacobian matrix of the piecewise-linear transformation is unique for a given simplex Δ and has the form

$$\bar{J}_h(q^0, q^1, \ldots, q^d) = \begin{pmatrix} V(q^0) & \cdots & V(q^d) \end{pmatrix} \begin{pmatrix} Tq^0 & \cdots & Tq^d \\ 1 & \cdots & 1 \end{pmatrix}^{-1} =: V_\Delta Z_\Delta^{-1}$$

with $V_\Delta \in \mathbb{R}^{d+1}$ and $Z_\Delta^{-1} \in \mathbb{N}^{(d+1)\times(d+1)}$.

Proof. It follows from Lemma 3.3.1 that there is a unique representation of points $q \in \Delta(q^0, q^1, \ldots, q^d)$ and $z = Tq$ given by

$$\begin{pmatrix} z \\ 1 \end{pmatrix} = \begin{pmatrix} Tq \\ 1 \end{pmatrix} = \begin{pmatrix} Tq^0 & \cdots & Tq^d \\ 1 & \cdots & 1 \end{pmatrix} \mu = Z_\Delta \mu \tag{3.22}$$

We now define the affine function $V_h(q)$, with $V_h(q^i) = V(q^i)$ for $i = 0, 1, \ldots, d$, as

$$V_h(q) = \begin{pmatrix} V(q^0) & \cdots & V(q^d) \end{pmatrix} \cdot \mu = V_\Delta Z_\Delta^{-1} \cdot \begin{pmatrix} z \\ 1 \end{pmatrix} = \bar{J}_h \cdot \begin{pmatrix} Tq \\ 1 \end{pmatrix} \tag{3.23}$$

Then, the claim follows from writing the affine function as $V_h(q) = \nabla V_\Delta \cdot q + V_0$. We note that both matrices Z_Δ and Z_Δ^{-1} belong to $\mathbb{N}^{(d+1)\times(d+1)}$ due to the fact that hypercubes in $\mathcal{S}_C$ have unit volume, i.e., $\det(Z_\Delta) = 1$. ■

We then show that force-stepping trajectories $\{t_k, q_k, \dot{q}_k\}$, as described in Definition 3.2.1, can be systematically tracked in phase space. In particular, we show that successive times of boundary crossings $t_0 < t_1 < t_2 < \ldots$ and the corresponding sequence of positions $q_0, q_1, q_2, \ldots$ can be obtained from the simplicial partition presented above. The procedure applied for solving this problem, namely *the time-step problem* (see Proposition 3.3.1), is indeed the method TIME-STEP required by the implementation of the force-stepping integrator presented in Algorithm 3,

$$\{t_{k+1}, B_s\} \leftarrow \text{TIME-STEP}\left(q_k, \dot{q}_k, \nabla V_k, t_k; Z_{\Delta_k}^{-1}\right)$$

Proposition 3.3.1 (The time-step problem) *A general trajectory of the form* $q(t) = q_k + (t - t_k)\dot{q}_k - \frac{1}{2}(t - t_k)^2 M^{-1}\nabla V_k$, *with* $q(t_k^+) \in \Delta(q^0, q^1, \ldots, q^d)$ *and* $t > t_k$, *intersects the boundary* B_s *of the simplex* $\Delta(q^0, q^1, \ldots, q^d)$ *at* $q(t_{k+1}) = q_{k+1}$ *with*

$$t_{k+1} = \min_{r\in[0,d]} \inf \left\{ t > t_k : \mu_{1,r} + (t - t_k)\mu_{2,r} + (t - t_k)^2\mu_{3,r} = 0 \right\}$$

$$s = \arg\min_{r\in[0,d]} \left\{ \mu_{1,r} + (t_{k+1} - t_k)\mu_{2,r} + (t_{k+1} - t_k)^2\mu_{3,r} = 0 \right\}$$

where

$$\left(\begin{array}{c|c|c} \mu_1 & \mu_2 & \mu_3 \end{array} \right) = Z_\Delta^{-1} \left(\begin{array}{c|c|c} Tq_k & T\dot{q}_k & -M^{-1}T\nabla V_k/2 \\ 1 & 0 & 0 \end{array} \right)$$

Proof. It follows from Definition 3.3.1 and Lemma 3.3.1 that a trajectory of the form $q(t) = \sum_{r=0}^{d} \mu_r(t) q^r$, with $q(t_k^+) \in \Delta(q^0, q^1, \ldots, q^d)$ and $t > t_k$, intersects the boundary B_{r^*} of the simplex $\Delta(q^0, q^1, \ldots, q^d)$ at $q(t^*)$ if and only if the r^*-th component of $\mu(t^*)$ is equal to zero with all other components positive. Then, the following equality holds

$$\mu(t) = Z_\Delta^{-1} \left(\begin{array}{c} Tq(t) \\ 1 \end{array} \right) = \underbrace{Z_\Delta^{-1} \left(\begin{array}{c} Tq_k \\ 1 \end{array} \right)}_{\mu_1 \in \mathbb{R}^{d+1}} + (t - t_k) \underbrace{Z_\Delta^{-1} \left(\begin{array}{c} T\dot{q}_k \\ 0 \end{array} \right)}_{\mu_2 \in \mathbb{R}^{d+1}} + (t - t_k)^2 \underbrace{Z_\Delta^{-1} \left(\begin{array}{c} -M^{-1}T\nabla V_k/2 \\ 0 \end{array} \right)}_{\mu_3 \in \mathbb{R}^{d+1}}$$

and the claim follows from the fact that t_{k+1} is the earliest $t^* > t_k$ for which $\mu_{r^*=s}(t^*) = 0$ among all possible pairs (t^*, r^*). ■

It is worth noting that the representation of the approximate potential energy V_h can be restricted solely to the current simplex Δ_k and the relevant simplex-related matrices are $\langle V_{\Delta_k}, Z_{\Delta_k}, Z_{\Delta_k}^{-1} \rangle$, with

$$V_{\Delta_k} = \left(\begin{array}{ccccccc} V(q^0) & \cdots & V(q^s) & \cdots & V(q^d) \end{array} \right) \tag{3.24a}$$

$$Z_{\Delta_k} = \left(\begin{array}{ccccc} z^0 & \cdots & z^s & \cdots & z^d \\ 1 & \cdots & 1 & \cdots & 1 \end{array} \right) \tag{3.24b}$$

We then claim that *the replacement rule* (see Chien and Kuh [11] and Lemma 3.3.3) provides for a very efficient scheme to construct the adjacent simplex Δ_{k+1} knowing only the boundary B_s of the current simplex Δ_k intersected by the trajectory. Moreover, we describe an efficient procedure for updating all simplex-related matrices required to compute the force-stepping trajectory in Δ_{k+1}. In particular, given all Δ_k-related matrices and the intersected boundary B_s, matrices are updated as

follows

$$V_{\Delta_{k+1}} = \begin{pmatrix} V(q^0) & \cdots & V(q'^s) & \cdots & V(q^d) \end{pmatrix} \tag{3.25a}$$

$$Z_{\Delta_{k+1}} = \begin{pmatrix} z^0 & \cdots & z'^s & \cdots & z^d \\ \\ 1 & \cdots & 1 & \cdots & 1 \end{pmatrix} \tag{3.25b}$$

$$Z^{-1}_{\Delta_{k+1}} = Z^{-1}_{\Delta_k} - \frac{\left(Z^{-1}_{\Delta_k} u^s\right) \otimes \left(v^s Z^{-1}_{\Delta_k}\right)}{1 + v^s Z^{-1}_{\Delta_k} u^s} \tag{3.25c}$$

where $z'^s = z^{s+1} + z^{s-1} - z^s$, $q'^s = T^{-1} z'^s$, $u^s = z'^s - z^s$, and $v^s_i = \delta_{i,s}$. Matrices $Z_{\Delta_{k+1}}$ and $V_{\Delta_{k+1}}$ are obtained by replacing z^s with z'^s and $V(q^s)$ with $V(q'^s)$, respectively. The matrix $Z^{-1}_{\Delta_{k+1}}$ is 1-rank updated by means of the Sherman-Morrison formula [28]—with algorithm complexity $O(d^2)$—instead of computing the inverse of $Z_{\Delta_{k+1}}$—with algorithm complexity $O(d^3)$. The overall gain in efficiency is remarkable.

Lemma 3.3.3 (The replacement rule) *Let B_s be the boundary shared by two adjacent simplices. Given the simplex $\Delta(q^0, q^1, \ldots, q^{s-1}, q^s, q^{s+1}, \ldots, q^d)$, with $Tq^0 \leqslant Tq^1 \leqslant \ldots \leqslant Tq^d$, its neighbor is simply defined by replacing q^s with q'^s. The new vertex is defined as*

$$q'^s = q^{s+1} + q^{s-1} - q^s$$

where $q^{d+1} \equiv q^0$ and $q^{-1} \equiv q^d$.

Corollary 3.3.2 *The replacement rule preserves the order in the set of vertices $\mathcal{V}$.*

Proof. For $s = 1, \ldots, d-1$, the following relation holds $Tq^{s-1} \leqslant Tq'^s \leqslant Tq^{s+1}$. For $s = 0$ ($s = d$) a simple backward (forward) shift of indices has to be introduced in the updated set of vertices. After the shift is performed, the new set of vertices verifies the ordering relation. ■

We conclude this section with the second method required by the implementation of the force-stepping integrator presented in Algorithm 3,

$$\left\{\nabla V_{k+1}, \langle V_{\Delta_{k+1}}, Z_{\Delta_{k+1}}, Z^{-1}_{\Delta_{k+1}}\rangle\right\} \leftarrow \textsc{Update}\left(\nabla V_k, \langle V_{\Delta_k}, Z_{\Delta_k}, Z^{-1}_{\Delta_k}\rangle; B_s\right)$$

The method is defined in Algorithm 4 and is responsible for applying *the replacement rule* as well as updating ∇V_k and all simplex-related matrices $\langle V_{\Delta_k}, Z_{\Delta_k}, Z_{\Delta_k}^{-1}\rangle$. We remark that, for all practical purposes, there is no need to shift indices after vertices q^0 or q^d are updated by the algorithm, as suggested by Corollary 3.3.2.

Algorithm 4 UPDATE$\big(\nabla V_k, \langle V_{\Delta_k}, Z_{\Delta_k}, Z_{\Delta_k}^{-1}\rangle; B_s\big)$

Require: $z'^s = z^{s+1} + z^{s-1} - z^s$, $u^s = z'^s - z^s$, $v_i^s = \delta_{i,s}$ and $q'^s = T^{-1}z'^s$

1: $Z_{\Delta_{k+1}}^{-1} \leftarrow Z_{\Delta_k}^{-1} - \frac{(Z_{\Delta_k}^{-1}u^s)\otimes(v^s Z_{\Delta_k}^{-1})}{1+v^s Z_{\Delta_k}^{-1}u^s}$

2: $Z_{\Delta_{k+1}} \leftarrow \begin{pmatrix} z^0 & \cdots & z'^s & \cdots & z^d \\ 1 & \cdots & 1 & \cdots & 1 \end{pmatrix}$

3: $V_{\Delta_{k+1}} \leftarrow \begin{pmatrix} V(q^0) & \cdots & V(q'^s) & \cdots & V(q^d) \end{pmatrix}$

4: $\bar{J}_h \leftarrow V_{\Delta_{k+1}} Z_{\Delta_{k+1}}^{-1}$

5: $\nabla V_{k+1} \leftarrow \bar{J}_{h,[1\ldots d]} T$

6: **return** $\nabla V_{k+1}, \langle V_{\Delta_{k+1}}, Z_{\Delta_{k+1}}, Z_{\Delta_{k+1}}^{-1}\rangle$

3.4 Conservation properties

Some symmetries of Lagrangian systems, such as time-reversal, parity and invariance under time shifts in the particular case of autonomous or time-independent systems, are preserved upon piecewise-linear approximation of the potential energy, which results in the exact conservation of the corresponding momentum maps. In particular, force-stepping trajectories conserve energy exactly. By contrast, and unlike energy-stepping, which results in exact conservation of all momenta, piecewise-linear approximation may break some symmetries of the potential energy, such as translation and rotation invariance. resulting in lack of exact conservation of the corresponding momenta. However, in this section we show that the force-stepping scheme is nearly-conserving for all symmetries of the system, whether explicitly known or not. The numerical experiments presented in Section 3.6 further show that the corresponding momentum maps tend to remain nearly constant uniformly for all times. We also show that some broken symmetries can be restored exactly by recourse to Lagrangian reduction.

3.4.1 Conservation of momentum maps

We recall (cf., e. g., [57]) that a general Lagrangian is a function $L : TQ \to \mathbb{R}$, where Q is a smooth manifold, known as configuration manifold, and TQ is the corresponding tangent bundle, consisting of pairs of configurations and velocities. For simplicity, we restrict attention to time-independent or autonomous Lagrangians. Let X denote some suitable topological space of trajectories $q : [0, T] \to Q$ joining fixed initial and final configurations $q(0)$ and $q(T)$, respectively. Then, the action integral $I : X \to \mathbb{R}$ over the time interval $[0, T]$ is

$$I(q) = \int_0^T L\big(q(t), \dot{q}(t)\big)\, dt \tag{3.26}$$

where we assume sufficient regularity of L and $q(t)$ for all mathematical operations to be well defined. According to Hamilton's principle, the physical trajectories of the system are the critical points of I, i.e., $q \in X$ is a trajectory if

$$\delta I(q, \varphi) = 0 \tag{3.27}$$

for all variations $\varphi \in C_c^\infty(0, T)$. Throughout the present work the configurational space of interest is $Q = E(n)^N$, where $E(n)$ is the Euclidean space of dimension n (i.e., $d = nN$), and the Lagrangian is assumed to be of the form (3.1), with V smooth and bounded below, or (3.2) with V_h continuous piecewise-linear. An appropriate space of trajectories is $X = W^{2,\infty}_{\mathrm{loc}}([0, \infty))$, as demonstrated in Section 3.5. We note that in the case of a continuous piecewise-linear potential energy the Euler-Lagrange equations are not defined in the classical sense. However, the trajectories can still be understood as critical points of the action functional I_h as shown in Theorem 3.2.1. In particular, on element boundaries the acceleration is not uniquely defined. However, we show in Section 3.5 that, for almost all initial conditions, simplex boundaries are crossed only at isolated points in time, and for those initial conditions the discrete trajectories are well defined and convergent.

Let G be a Lie group with Lie algebra $\mathfrak{g} = T_eG$. A left action of G on Q is a mapping $\Phi : G \times Q \to Q$ such that: i) $\Phi(e, \cdot) = \mathrm{id}$; ii) $\Phi(g, \Phi(h, \cdot)) = \Phi(gh, \cdot)\ \forall g, h \in G$. Let $\xi \in \mathfrak{g}$. Then, the

infinitesimal generator of Φ corresponding to ξ is the vector field $\xi_Q \in TQ$ given by

$$\xi_Q(q) = \frac{d}{dt}\left[\Phi(\exp(t\xi), q)\right]_{t=0} \tag{3.28}$$

The momentum map $J : TQ \to \mathfrak{g}^*$ defined by the action Φ then follows from the identity

$$\langle J(q,\dot{q}), \xi\rangle = \langle \partial_{\dot{q}} L(q,\dot{q}), \xi_Q(q)\rangle, \quad \forall \xi \in \mathfrak{g} \tag{3.29}$$

We say that the Lagrangian L is invariant under the action Φ if

$$L(\Phi_g(q), T\Phi_g(q)\dot{q}) = L(q,\dot{q}), \quad \forall g \in G,\ (q,\dot{q}) \in TQ \tag{3.30}$$

where we write $\Phi_g(\cdot) = \Phi(g,\cdot)$. Under these conditions, we additionally say that G is a symmetry group of the system and that Φ expresses a symmetry of the system.

The classical theorem of Noether states that if L is invariant under the action Φ, then the corresponding momentum map J is a constant of the motion, i.e., it remains constant along trajectories. Classical examples include:

i) *Conservation of linear momentum.* In this case, $Q = E(n)^N$, $G = E(n)$ and $\Phi(u,q) = \{q_1+u, \ldots, q_N+u\}$ represent a rigid translation of the system by $u \in E(n)$. The corresponding momentum map is the total linear momentum of the system, $J = p_1 + \cdots + p_N$. If the Lagrangian is invariant under translations, then the total linear momentum is a constant of the motion.

ii) *Conservation of angular momentum.* In this case, $Q = E(n)^N$, $G = SO(n)$ and $\Phi(R,q) = \{Rq_1, \ldots, Rq_N\}$ represent a rigid rotation of the system by $R \in SO(n)$. The corresponding momentum map is the total angular momentum of the system, $J = q_1 \times p_1 + \cdots + q_N \times p_N$. If the Lagrangian is invariant under rotations, then the total angular momentum is a constant of the motion.

Conservation of energy can be fit into this framework by recourse to a space-time formulation in which time is regarded as a generalized coordinate, e. g., q_0. The corresponding space-time Lagrangian is $\mathbb{L}(\mathbb{q}, \mathbb{q}') = \mathbb{L}((q_0, q), (q_0', q')) = L(q, q'/q_0', q_0)$, where $L(q, \dot{q}, t)$ is a general time-dependent Lagrangian. The space-time configuration manifold is $\mathbb{Q} = \mathbb{R} \times Q$. Let $G = \mathbb{R}$, $\Phi(s, q) = (q_0 + s, q)$ is a time-shift and suppose that $\mathbb{L}$ is invariant under Φ, i.e., L is time-independent. Then $J = L - \partial_{\dot{q}} L \cdot \dot{q} = -E$, i.e., the total energy of the system, is a constant of the motion. However, exact energy conservation follows more directly as a consequence of (3.11), i.e., energy conservation is built explicitly into the force-stepping scheme.

Evidently, a naïve continuous piecewise-linear approximation of the potential energy breaks the symmetries of the system in general, including translation and rotation invariance. Thus, if G is a symmetry group of V and Φ is an action that leaves V invariant, the corresponding momentum map J may not be constant along force-stepping trajectories in general. This lack of exact conservation can be remedied by recourse to Lagrangian reduction, as shown subsequently. However, even the naïve force-stepping scheme, with no symmetrization of the approximated potential, has near-conservation properties, as shown next.

Theorem 3.4.1 *Suppose that G is a symmetry group of the potential V, i.e., $V \circ \Phi_g = V$ for all $g \in G$, and let $J(q, \dot{q})$ be the corresponding momentum map. Let $q_0, \dot{q}_0 \in \mathbb{R}^{nN}$ and $h_m \to 0$. Suppose, in addition, that $q_{h_m} \in W^{2,\infty}(0, T)$ are force-stepping trajectories corresponding to the approximate potentials V_{h_m} and q is a trajectory corresponding to the original potential V, with $q(0) = q_{h_m}(0) = q_0$ and $\dot{q}(0) = \dot{q}_{h_m}(0) = \dot{q}_0$ for all m. Then we have $J(q_{h_m}, \dot{q}_{h_m}) \to J(q, \dot{q}) \equiv J(q_0, \dot{q}_0)$ in $W^{1,\infty}_{\mathrm{loc}}([0, \infty))$ for almost all initial conditions $(q_0, \dot{q}_0)$.*

Proof. From Theorem 3.5.2 we deduce that, except for a $(2nN - 1)$-dimensional set of initial conditions $(q_0, \dot{q}_0)$, q_{h_m} converges to the continuum solution q strongly in $W^{2,\infty}_{\mathrm{loc}}([0, \infty))$. Since L and, thus, J are smooth this implies that $J(q_{h_m}, \dot{q}_{h_m}) \to J(q, \dot{q})$ in $W^{1,\infty}_{\mathrm{loc}}([0, \infty))$. As J is conserved along the continuum trajectory $(q, \dot{q})$, i.e., $J(q, \dot{q}) \equiv J(q_0, \dot{q}_0)$, the claim follows. ■

It is interesting to note that an *approximate Noether theorem* can also be obtained for general potentials, not just those resulting by approximation of a symmetric potential. We state it here for

smooth potentials and compact *almost-symmetry groups*. In order to quantify the extent to which a Lagrangian L fails to be symmetric with respect to a Lie group G acting on Q via Φ, we introduce the functions

$$L_g(q, \dot{q}) = L(\Phi_g(q), T\Phi_g(q)\dot{q}) \tag{3.31}$$

parameterized by $g \in G$. In the symmetric case, $L_g = L_e = L$ exactly. The following theorem shows that, for an approximate conservation estimate, we require that $L_g \sim L$ up to first order in q and up to second order in $\dot{q}$. In particular, these assumptions are satisfied for the piecewise-linear Lagrangians (3.2), which approximate the original Lagrangian (3.1) to first order in q and to second order in $\dot{q}$.

Theorem 3.4.2 (Approximate Noether's theorem) *Suppose G is a compact Lie group and Φ is a smooth left action of G on Q with momentum map J. Let $U \subset TQ$ be relatively compact and assume that L is uniformly strictly convex on U with respect to $\dot{q}$. Suppose (near symmetry) that*

$$\sup_{g \in G} \left(\|L - L_g\|_{L^\infty(U)} + \left\| \frac{\partial L}{\partial q} - \frac{\partial L_g}{\partial q} \right\|_{L^\infty(U)} + \left\| \frac{\partial L}{\partial \dot{q}} - \frac{\partial L_g}{\partial \dot{q}} \right\|_{L^\infty(U)} + \left\| \nabla^2_{\dot{q}} L - \nabla^2_{\dot{q}} L_g \right\|_{L^\infty(U)} \right) \leq \varepsilon \tag{3.32}$$

Then, for all trajectories $(q, \dot{q})$ such that $(q(t), \dot{q}(t)) \in \overline{U}$ for $t \in [0, T]$, there exists a constant $C > 0$ such that

$$\|J(q, \dot{q}) - J(q_0, \dot{q}_0)\|_{W^{1,\infty}([0,T])} \leq C\varepsilon \tag{3.33}$$

Note that for Lagrangians of the form (3.1), for given T, q_0 and $\dot{q}_0$ one can find a relatively compact set U satsifying the assumptions of Theorem 3.4.2 which only depends on the minimum value of V. This is easily deduced from energy conservation.

Proof. Define the symmetrized Lagrangian L' by

$$L'(q, \dot{q}) := \fint_G L(\Phi_g(q), T\Phi_g(q)\dot{q})\, dg \tag{3.34}$$

where $\fint_G (\cdots)\, dg = \frac{1}{|G|} \int_G (\cdots)\, dg$ denotes the normalized integration with respect to the right

invariant Haar measure of G. Then L' is invariant under Φ since

$$L'(\Phi_{\tilde{g}}(q), T\Phi_{\tilde{g}}(q)\dot{q}) = \fint_G L(\Phi_g(\Phi_{\tilde{g}}(q)), T\Phi_g(\Phi_{\tilde{g}}(q))T\Phi_{\tilde{g}}(q)\dot{q})\, dg = \fint_G L(\Phi_{g\tilde{g}}(q), T\Phi_{g\tilde{g}}(q)\dot{q})\, dg = \fint_G L(\Phi_g(q), T\Phi_g(q)\dot{q})\, dg = L'(q,\dot{q}) \quad (3.35)$$

for all $\tilde{g} \in G$. By Noether's Theorem we thus conclude that the momentum map J' corresponding to L' is conserved along trajectories q' corresponding to the Lagrangian L', i.e., on paths satisfying

$$\frac{\partial L'}{\partial q}(q'(t)) - \frac{d}{dt}\frac{\partial L'}{\partial \dot{q}}(q'(t)) = 0 \quad (3.36)$$

Now we have, by assumption,

$$\|L - L'\|_{L^\infty(U)} = \left\| \fint_G (L - L_g)\, dg \right\|_{L^\infty(U)} \leq \fint_G \|L - L_g\|_{L^\infty(U)}\, dg \leq \varepsilon \quad (3.37)$$

since $\fint_G 1\, dg = 1$. In addition,

$$\left\| \frac{\partial L}{\partial q} - \frac{\partial L'}{\partial q} \right\|_{L^\infty(U)} = \left\| \fint_G \left(\frac{\partial L}{\partial q} - \frac{\partial L_g}{\partial q} \right) dg \right\|_{L^\infty(U)} \leq \fint_G \left\| \frac{\partial L}{\partial q} - \frac{\partial L_g}{\partial q} \right\|_{L^\infty(U)} dg \leq \varepsilon \quad (3.38)$$

and, similarly,

$$\left\| \frac{\partial L}{\partial \dot{q}} - \frac{\partial L'}{\partial \dot{q}} \right\|_{L^\infty(U)} \leq \varepsilon \quad \text{and} \quad \left\| \nabla^2_{\dot{q}} L - \nabla^2_{\dot{q}} L' \right\|_{L^\infty(U)} \leq \varepsilon \quad (3.39)$$

Also by assumption, the Hessian $\nabla^2_{\dot{q}} L$ is uniformly non-singular. The previous estimate (3.39) shows that this is also true for $\nabla^2_{\dot{q}} L'$, as we may, without loss of generality, assume that ε is sufficiently small. As a consequence, we obtain that the solutions q and q' of

$$\frac{\partial L}{\partial q}(q(t)) - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}(q(t)) = 0 \quad \text{resp.} \quad \frac{\partial L'}{\partial q}(q'(t)) - \frac{d}{dt}\frac{\partial L'}{\partial \dot{q}}(q'(t)) = 0 \quad (3.40)$$

subject to the same initial conditions $q(0) = q'(0) = q_0$, $\dot{q}(0) = \dot{q}'(0) = \dot{q}_0$ satisfy the estimate

$$\|q - q'\|_{W^{2,\infty}(0,T)} \leq C\varepsilon \tag{3.41}$$

Another consequence of the above estimates (3.37), (3.38) and (3.39) is that, since by (3.29) the momentum maps J and J' depend smoothly on $\partial_{\dot{q}} L$ and $\partial_{\dot{q}} L'$, respectively,

$$\|J - J'\|_{W^{1,\infty}(U)} \leq \|L - L'\|_{W^{1,\infty}(U)} \leq C\varepsilon \tag{3.42}$$

(Note that $W^{1,\infty}$ is an algebra.) Finally, from (3.41) and (3.42) it follows that

$$\|J(q,\dot{q}) - J(q_0,\dot{q}_0)\|_{W^{1,\infty}(0,T)} \leq$$
$$\|J(q,\dot{q}) - J'(q',\dot{q}')\|_{W^{1,\infty}(0,T)} + \|J'(q_0,\dot{q}_0) - J(q_0,\dot{q}_0)\|_{W^{1,\infty}(0,T)} \leq C\varepsilon \tag{3.43}$$

■

In essence, the preceding theorem states that almost-symmetries, in the sense of (3.32), result in almost-conservation in the sense of (3.33). The theorem thus effectively removes the rigidity of the classical Noether's theorem, which applies to exact symmetries only.

3.4.2 Lagrangian reduction

A general strategy for avoiding broken symmetries in the force-stepping scheme is supplied by the theory of Lagrangian reduction. The aim of this theory is to reduce the dimension of the configuration space of a Lagrangian by systematically exploiting constants of motion and symmetry groups (cf., e. g., [57, 59] and references therein). Thus, suppose G is a Lie group, $L : TQ \to \mathbb{R}$ is a Lagrangian invariant under a smooth left action Φ of G on Q, with momentum map J. Then, the theory defines a reduced Lagrangian $l : TQ/G \to \mathbb{R}$ in which the group coordinates are eliminated. Classical examples include:

i) *Total linear momentum.* In this case, $Q = E(n)^N$, $G = E(n)$ and $J = p_1 + \cdots + p_N$. Then, given a total linear momentum $\mu \in \mathfrak{g}^*$, the isotropy group G_μ, i.e., the group of actions that leaves the level set $J^{-1}(\mu)$ invariant, reduces to $G_\mu = G$. Thus, the corresponding reduced phase space is given by $J^{-1}(\mu)/G_\mu = T(E(n)^{N-1})$ and has dimension $2n(N-1)$.

ii) *Total angular momentum ($n = 3$, $N \geq 3$).* In this case, $Q = E(3)^N$, $G = SO(3)$ and $J = q_1 \times p_1 + \cdots + q_N \times p_N$. Then, given a total angular momentum $\mu \in \mathfrak{g}^*$, the dimension of the corresponding reduced phase space $J^{-1}(\mu)/G_\mu$ depends on the value of μ. For $\mu = 0$, any rotation leaves the level set $J^{-1}(0)$ invariant, that is $G_\mu = G$, and therefore the reduced phase space has dimension $6(N-1)$. For $\mu \neq 0$, only rotations along the direction of μ leave the level set $J^{-1}(\mu)$ invariant, that is $G_\mu = S^1$, and therefore the reduced phase space has dimension $6N - 4$.

iii) *Total linear and angular momenta ($n = 3$, $N \geq 3$).* Given a total linear momentum μ_1 and a total angular momentum μ_2, one concludes from the preceding analysis that the corresponding reduced phase space TQ/G has dimension $6N - 12$, resp. $6N - 10$, for $\mu_2 = 0$, resp. $\mu_2 \neq 0$.

The reduced Lagrangian resulting from the elimination of constrained degrees of freedom represents the intrinsic or internal dynamics of the system. Furthermore, the reduced Lagrangian implicitly inherits the conservation laws of the original system. In addition, it is possible to reconstruct the original dynamics from the reduced dynamics.

As an application of the theory, consider Lagrangians of the form (3.1) with diagonal mass matrix M and potential V, invariant under translations and rotations. Such systems are known as N-body problems and are the subject of a vast body of literature (cf., e. g., [54,55,60,69,81]). Reduction with respect to translational symmetry may be achieved by recourse to Jacobi coordinates $q' \in E(n)^{N-1}$ and center of mass coordinates $q_{cm} \in E(n)$, defined as

$$q'_j = q_{j+1} - \frac{\sum_{i=1}^{j} m_i q_i}{\sum_{i=1}^{j} m_i}, \quad 1 \leq j < N \tag{3.44}$$

$$q_{cm} = \frac{\sum_{i=1}^{N} m_i q_i}{m_T}, \tag{3.45}$$

where m_i is the mass of the i-th particle and

$$m_T = \sum_{i=1}^{N} m_i \tag{3.46}$$

is the total mass of the system. Clearly, the map by the map $\pi : q \rightarrow (q', q_{cm})$ is linear and bijective, and its inverse $\pi^{-1}(q', q_{cm}) \rightarrow q$ is given by

$$q_j = q_{cm} + q'_{j-1}\frac{\sum_{i=1}^{j-1} m_i}{\sum_{i=1}^{j} m_i} - \sum_{i=j+1}^{N} \frac{m_i q'_{i-1}}{\sum_{k=1}^{i} m_k}, \quad 1 \leq j \leq N \tag{3.47}$$

with $q'_0 = 0$. The corresponding constant of the motion is the total linear momentum $\mu_1 = \sum_{i=1}^{N} m_i \dot{q}_i$ and the reduced Lagrangian

$$l(q', \dot{q}') = \frac{1}{2}\dot{q}'^T M' \dot{q}' - V'(q') \tag{3.48}$$

retains the desirable form (3.1). It is readily shown that the reduced mass matrix M' is diagonal with nonzero elements

$$m'_j = m_{j+1}\frac{\sum_{i=1}^{j} m_i}{\sum_{i=1}^{j+1} m_i} \tag{3.49}$$

and that the reduced potential is

$$V'(q') = V\left(\pi^{-1}(q', 0)\right) \tag{3.50}$$

It should be carefully noted that the evolution of the center of mass $q_{cm}(t)$ is decoupled from the dynamics of the reduced system and, as part of the reconstruction of the original dynamics, it can be trivially obtained as

$$\dot{q}_{cm} = \frac{\mu_1}{m_T} \tag{3.51}$$

The trajectories of the reduced Lagrangian system can now be approximated by replacing (3.48) by an approximate reduced Lagrangian $l_h(q', \dot{q}')$ that can be solved exactly, e. g. of the form

$$l_h(q', \dot{q}') = \frac{1}{2}\dot{q}'^T M' \dot{q}' - V'_h(q') \tag{3.52}$$

obtained by replacing V' by its piecewise-linear approximation V'_h. By this construction, the approximate Lagrangian l_h is translation invariant and the total linear momentum is exactly conserved along the corresponding trajectories.

The reduction with respect to the rotational symmetry, in applications which involve an arbitrary number of particles ($N \geq 3$) and non-vanishing angular momentum, is presently the subject of active research (cf., e. g., [54, 55, 69, 81]). Unfortunately, it is not possible to construct internal or shape coordinates which result in a reduced Lagrangian of the form (3.1).

3.4.3 Conservation of discrete symmetries

Lagrangian systems may also have some symmetries that are associated with actions of *discrete* groups rather than with actions of *continuous* groups—Lie groups G presented in previous subsections are continuous. This is the case of many non-relativistic systems which are invariant under parity. Parity is a discrete symmetry expressed by an operator P, whose action on configuration space $Q = E(3)^N$ is given by $Pq_j = -q_j$, and on $Q = E(2)^N$ is given by flipping the sign of only one component of q_j. In general, G-invariant reductions may or may not preserve the parity symmetry P. However, reductions with respect to translations given by the map $\pi(q)$, (3.44)-(3.45), induce effective potentials $V'(q')$ invariant under parity, that is $V'(Pq') = V'(q')$.

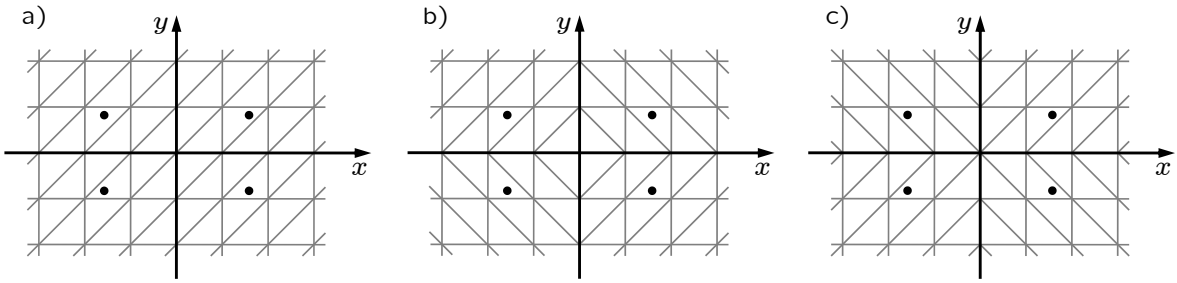


Figure 3.3: Three feasible regular triangulations of $\mathbb{R}^2$. Triangulation a) breaks the parity invariance of $V(q)$; triangulations b) and c), do not.

In applications, we employ regular triangulations $\mathcal{T}_h$ which result in piecewise-affine approximate potentials V'_h invariant under parity. Figure 3.3 illustrates the simple case of $Q' = E(2)$. The extension to $Q' = E(n)^{N-1}$ is straightforward and falls squarely within the construction presented

in Section 3.3. In addition, time-reversibility follows directly from the definition of the force-stepping scheme. The time-reversal symmetry of Lagrangian systems is a discrete symmetry that is widely considered to be essential to the design of efficient time-integration schemes (cf., e. g., [30] and references therein).

3.4.4 Conservation of the symplectic structure

In order to establish the symplecticity of the force-stepping scheme, we proceed to look at the Lagrangian system defined by (3.1) from a Hamiltonian perspective. To this end, we introduce the phase space $P = T^*Q$, consisting of pairs (q, p) of configurations $q \in Q = E(n)^N$ and momenta $p \in T_q^*Q$, and the Hamiltonian $H : P \to \mathbb{R}$ as

$$H(q,p) = \sup_{v \in T_qQ} \{p \cdot v - L(q,v)\} = \frac{1}{2} p^T M^{-1} p + V(q) \tag{3.53}$$

For a G-invariant reduced Lagrangian defined by (3.48), we also introduce the reduced Hamiltonian $H_\mu : P_\mu \to \mathbb{R}$ as

$$H_\mu(q',p') = \sup_{v \in T_{q'}Q/G} \{p' \cdot v - l(q',v)\} = \frac{1}{2} p'^T M'^{-1} p' + V'(q') \tag{3.54}$$

Likewise, the Hamiltonian corresponding to the approximate reduced Lagrangian (3.52) is

$$H_{\mu,h}(q',p') = \sup_{v \in T_{q'}Q/G} \{p' \cdot v - l_h(q',v)\} = \frac{1}{2} p'^T M'^{-1} p' + V'_h(q') \tag{3.55}$$

We now endow P with the canonical *symplectic two-form*

$$\Omega = dq_1 \wedge dp_1 + \cdots + dq_N \wedge dp_N \tag{3.56}$$

Then, pairs (P, Ω) and (P_μ, Ω_μ) define symplectic manifolds, where the reduced symplectic form Ω_μ is determined by the *Symplectic Reduced Theorem* [59]. For systems with Abelian symmetries, Ω_μ

carries the canonical symplectic structure modified by magnetic terms β_μ. In particular, reductions with respect to translations given by the map $\pi(q)$, (3.44)-(3.45), induce the following *reduced symplectic two-form*

$$\Omega_{\mu_1} = \Omega - \beta_{\mu_1} = q_1' \wedge dp_1' + \cdots + dq_{N-1}' \wedge dp_{N-1}' \tag{3.57}$$

where $p_i' = m_i' \dot{q}_i'$ and β_{μ_1} is the curvature of the mechanical connection, i.e., $\beta_{\mu_1} = dR \wedge d\mu_1$.

We recall (cf. e. g., [1, 2, 32]) that a diffeomorphism $\varphi : P \to P$ is *symplectic* if it preserves the symplectic two-form, i.e., if

$$\Omega(T\varphi(z)\eta_1, T\varphi(z)\eta_2) = \Omega(\eta_1, \eta_2) \tag{3.58}$$

Likewise, in reduced phase space, a diffeomorphism $\varphi : P_\mu \to P_\mu$ is symplectic if it preserves the reduced symplectic two-form. The symplecticity of Lipschitz homeomorphisms has been investigated by Whitney [79] and by Gol'dshtein and Dubrovskiy [21]. Then, in the present setting, it is possible to verify the symplecticity of force-stepping directly. To this end, we may write $\varphi(q_0', p_0') = (q(t), p(t))$ and

$$T\varphi \equiv \begin{pmatrix} \partial_{q'}\varphi_{q'} & \partial_{p'}\varphi_{q'} \\ \partial_{q'}\varphi_{p'} & \partial_{p'}\varphi_{p'} \end{pmatrix} = \begin{pmatrix} Q_{q'} & Q_{p'} \\ P_{q'} & P_{p'} \end{pmatrix} \tag{3.59}$$

then φ is symplectic if

$$P_{p'}^T Q_{p'} = Q_{p'}^T P_{p'} \tag{3.60}$$

$$P_{p'}^T Q_{q'} = Q_{p'}^T P_{q'} + I \tag{3.61}$$

$$Q_{q'}^T P_{q'} = P_{q'}^T Q_{q'} \tag{3.62}$$

We proceed to verify that these identities are indeed identically satisfied by the force-stepping scheme. For simplicity of notation and without loss of generality we take the mass matrix to be of the form $M = mI$. In addition, it suffices to consider mappings defined by the trajectory depicted in Figure 3.1, the general result then following by recursion. Evidently, the symplectic form is trivially conserved for $t \in [t_0, t_1)$, where t_1 is the time of intersection with the boundary of the underlying

triangulation. For $t \in (t_1, t_2)$, the dependence of t_1 on the initial conditions $(\dot{q}_0, q_0)$ must be carefully accounted for, i.e.,

$$\frac{\partial t_1}{\partial \dot{q}_0} = (t_1 - t_0)\frac{\partial t_1}{\partial q_0} = -\frac{t_1 - t_0}{\dot{q}(t_1)\cdot n} n = -\left|\frac{\partial t_1}{\partial \dot{q}_0}\right| n \tag{3.63}$$

For $t \in (t_1, t_2)$ we have

$$\dot{q}(q_0, \dot{q}_0, t) = \dot{q}_0 - \frac{t_1 - t_0}{m}\nabla V_I - \frac{t - t_1}{m}\nabla V_{II} \tag{3.64}$$

$$q(q_0, \dot{q}_0, t) = q_0 + (t - t_0)\dot{q}_0 - \frac{(2t - t_1 - t_0)(t_1 - t_0)}{2m}\nabla V_I - \frac{(t - t_1)^2}{2m}\nabla V_{II} \tag{3.65}$$

From these identities, the components of the Jacobian matrix $T\varphi$ are found to be

$$P_p = I - \frac{1}{m}\left|\nabla V_{II} - \nabla V_I\right|\left|\frac{\partial t_1}{\partial \dot{q}_0}\right| n \otimes n \tag{3.66}$$

$$Q_p = (t - t_0)I - \frac{t - t_1}{m}\left|\nabla V_{II} - \nabla V_I\right|\left|\frac{\partial t_1}{\partial \dot{q}_0}\right| n \otimes n \tag{3.67}$$

$$P_q = -\frac{1}{m}\left|\nabla V_{II} - \nabla V_I\right|\left|\frac{\partial t_1}{\partial q_0}\right| n \otimes n \tag{3.68}$$

$$Q_q = I - \frac{t - t_1}{m}\left|\nabla V_{II} - \nabla V_I\right|\left|\frac{\partial t_1}{\partial q_0}\right| n \otimes n \tag{3.69}$$

where the following identity has been used

$$\nabla V_{II} - \nabla V_I = \pm\left|\nabla V_{II} - \nabla V_I\right| n \tag{3.70}$$

A straightforward calculation shows that the symplecticity relations (3.60) are identically satisfied by (3.66), (3.67), (3.68) and (3.69), which establishes the the symplecticity of the force-stepping scheme.

3.4.5 Summary of the conservation properties of the force-stepping scheme

The results proven in the foregoing can be summarized as follows:

The force-stepping time-integration scheme is a symplectic, energy conserving, time-reversible integrator with automatic selection of the time step size. The scheme also conserves approximately all the momentum maps associated with the symmetries of the system. Exact conservation of momentum maps may be achieved by recourse to Lagrangian reduction.

Time-reversibility and parity-invariance of force-stepping follow directly from the definition of the scheme. The automatic time-step selection property also follows by construction. In particular, in regions where the velocity is high, the times effectively spent by the trajectory inside a simplex are short and the resulting time steps are small. By contrast, if the velocity is low, the resulting time steps are comparatively large.

Theorems 3.5.2 and 3.4.2 show that the naïve force-stepping scheme, while not exactly conserving in general, results in approximate momentum conservation, with the conservation error controlled uniformly on compact time intervals by the asymmetry of the approximate Lagrangian. This near-conservation property is frequently sufficient in applications and is born out by the numerical examples presented in Section 3.6.

We again emphasize that the momentum maps that are nearly conserved by force-stepping are precisely those conserved by the original system. This is in contrast to variational integrators, which conserve discrete momentum maps that are a discretization of—and differ from in general—the momentum maps of the original system.

3.5 Convergence analysis

Assume that the configuration space is $\mathbb{R}^d$ and the potential $V : \mathbb{R}^d \to \mathbb{R}$ is a C^1 function bounded from below. We use a regular triangulation $\mathcal{T}_h$ of $\mathbb{R}^d$ (see Section 3.3) and define V_h as the corresponding continuous piecewise affine approximation of V. Clearly, $V_h \to V$ uniformly on compact subsets of $\mathbb{R}^d$.

The approximating trajectories q_h can be represented by their successive times of element bound-

ary crossings $0 = t_0 < t_1 < t_2 < \ldots$ and the corresponding sequences of positions $q_h(0), q_h(t_1), q_h(t_2), \ldots$ and velocities $\dot{q}_h(0), \dot{q}_h(t_1), \dot{q}_h(t_2), \ldots$. Note that the trajectories for the piecewise-linear continuous approximation are defined unambiguously as long as they are never tangential to the boundary of any cell of the underlying triangulation while crossing this boundary. We will call trajectories such that, for every k, $\dot{q}_h(t_k)$ is not aligned with F, where F is a face of a triangulation element such that $q_h(t_k) \in F$, *transversal*. If q is non-transversal we denote by $t_{\max}$ the first time t for which $q(t)$ lies on an element boundary and $\dot{q}(t)$ is aligned with that boundary.

We also have to make sure that our iterative procedure defines an approximate trajectory for all positive times. This does not follow from transversality as the sequence (t_k) could be bounded. If this is indeed the case, we define $t_{\max} := \lim_{k\to\infty} t_k \in \mathbb{R}$, and our approximate trajectory will be defined only on $[0, t_{\max})$. For transversal trajectories with $t_k \to \infty$ as $k \to \infty$ we set $t_{\max} = \infty$.

The main goal of this section will be to justify our approximation scheme by proving that—except for a negligibly small set of initial conditions—the approximating trajectories converge to the original trajectory in a rather strong sense along any sequence of triangulation parameters h tending to zero. More precisely, the exceptional set is shown to be of Hausdorff dimension at most $2d-1$ in the $2d$-dimensional phase space (and consequently its Lebesgue measure vanishes).

By construction, an approximating trajectory q_h lies in $C^1([0, t_{\max}))$ and satisfies

$$M\ddot{q} = -\nabla V_h(q) \qquad \text{on} \qquad (0, t_{\max}) \tag{3.71}$$

for all times t such that $q(t)$ lies in the interior of some triangulation element. As an immediate consequence we obtain that q_h conserves energy:

Lemma 3.5.1 *If q_h is an approximating trajectory with initial conditions $(q_0, \dot{q}_0)$, then*

$$\frac{1}{2}\dot{q}_h^T(t) M \dot{q}_h(t) + V_h(q_h(t)) = \frac{1}{2}\dot{q}_0^T M \dot{q}_0 + V_h(q_0)$$

for all $t \in [0, t_{\max})$.

Proof. If $t \in [t_k, t_{k+1}]$, then

$$\frac{1}{2}\dot{q}_h^T(t)M\dot{q}_h(t) + V_h(q_h(t)) = \frac{1}{2}\dot{q}_h^T(t_k)M\dot{q}_h(t_k) + V_h(q_h(t_k)) \tag{3.72}$$

since q_h is a solution of the Euler-Lagrange equation (3.71) on $[t_k, t_{k+1}]$. The claim now follows by induction on n. ■

The approximating trajectories, being smooth inside the cells (a quadratic function) and C^1 across element boundaries, in fact belong to the Sobolev class $W^{2,\infty}_{\text{loc}}([0, t_{\max}))$ of twice weakly differentiable functions whose second derivative is bounded on any compact time interval. More precisely, we have the following

Lemma 3.5.2 *Approximating trajectories q_h are elements of $W^{2,\infty}_{\text{loc}}([0,\infty))$ if $t_{\max} = \infty$. If $t_{\max} < \infty$, then $q_h \in W^{2,\infty}([0, t_{\max}])$. Moreover, if $(q_h)_{h>0}$ is a family of approximating trajectories with the same initial conditions, then, for each finite $T \leq t_{\max}$, $\|q_h\|_{W^{2,\infty}([0,T])}$ is equi-bounded in h.*

Proof. Indeed, by Lemma 3.5.1 and the fact that V, and hence V_h, is bounded from below, $\dot{q}_h(t)$ is bounded uniformly in h and t and therefore $q_h(t)$ is bounded uniformly in h and uniformly in t on finite time intervals. But then also $\ddot{q}_h$ is bounded uniformly on finite time intervals as q_h satisfies (3.71) for almost all $t \in [0, t_{\max})$. In particular, if $t_{\max} < \infty$, then $\dot{q}_h$ can be extended as a Lipschitz function to $[0, t_{\max}]$. ■

For purposes of analysis, it is useful to note that in fact any non-transversal trajectory satisfies the Euler-Lagrange equation in the weak sense:

Proposition 3.5.1 *Suppose $t \mapsto q_h(t)$ is an approximating trajectory. Then* (3.71) *is satisfied in the weak sense, i.e., $\ddot{q}$ is the (piecewise continuous) weak second derivative of q and the equality is understood as equality almost everywhere on $(0, t_{\max})$.*

Proof. This follows from the proof of Lemma 3.5.2: Since q_h lies in $C^1([(0, t_{\max})), \ddot{q}$ is piecewise continuous with only finitely many jumps on any compact interval in $(0, t_{\max})$ and q_h satisfies (3.71) for all times $t \notin \{t_0, t_1, t_2, \ldots\}$, (3.71) is easily seen to hold in the sense of distributions. ■

Restricting our attention to transversal trajectories is justified by the following result.

Lemma 3.5.3 *Let $h > 0$. The set of initial conditions $(q_0, \dot{q}_0)$ for which the trajectory $(q_h(t), \dot{q}_h(t))$ is non-transversal has Hausdorff-dimension $2d - 1$.*

Proof. We first consider a single step $(q_{k-1}, \dot{q}_{k-1}) \to (q_k, \dot{q}_k)$ of the dynamics: Let $\partial\mathcal{T}_h$ be the collection of element boundaries and define the mapping $\tau : \partial\mathcal{T}_h \times \mathbb{R}^d \to \partial\mathcal{T}_h \times \mathbb{R}^d$ in the following way: For $(q, \dot{q}) \in \partial\mathcal{T}_h \times \mathbb{R}^d$ solve the Euler-Lagrange equation (3.71) backwards in time with initial condition $(q, \dot{q})$ and define $\tau(q, \dot{q})$ to be the position and velocity at the first element boundary crossing. This is a well defined mapping as long as $\dot{q}$ is not aligned with the boundary q lies on. In the exceptional case that $\dot{q}$ is aligned with this boundary we will view τ as being multivalued, more precisely $\tau(q, \dot{q})$ consisting of the set of transversal points in phase space $(\tilde{q}, \dot{\tilde{q}})$ that are mapped to $(q, \dot{q})$ under the discrete dynamics. (Note that $\#\tau(q, \dot{q})$ is bounded by the number of elements incident to $(q, \dot{q})$.)

Fix $T > 0$. Then on the set of those $(q, \dot{q})$ for which $\tau(q, \dot{q})$ is reached in a time span less than or equal to T, the mapping τ is locally Lipschitz. Since the set of non-transversal points $(q, \dot{q})$, i.e., for which q lies on some element boundary and $\dot{q}$ is parallel to this boundary, is $(2d - 2)$-dimensional, this proves that, for fixed $k \in \mathbb{N}$, the set of points $(q_1, \dot{q}_1)$ for which the corresponding trajectories satisfy $t_k \leq T$ and q_h is non-transversal at t_k is locally of finite $(2d - 2)$-dimensional Hausdorff measure.

Now note that clearly the pre-images of the mapping $(q_0, \dot{q}_0) \mapsto (q_1, \dot{q}_1)$ are of finite one-dimensional Hausdorff measure on the set of those $(q_0, \dot{q}_0)$ for which $t_1 \leq T$. Now finally sending $k \to \infty$ and $T \to \infty$, we obtain that the set of initial conditions for which t_k is non-transversal for some k is $(2d - 1)$-dimensional. ∎

We also need to show that the set of trajectories with $t_{\max} < \infty$ is negligible in a suitable sense. Note first that our approximating trajectories—being elements of $W^{2,\infty}(0, t_{\max})$—can be extended to functions in $C^1([0, t_{\max}])$. To this end, we introduce the following two subsets of $\mathbb{R}^d$: By Δ_1 denote the set of all vectors in $\mathbb{R}^d$ which are aligned with some triangulation element face. Since we use regular triangulations (see Section 3.3), Δ_1 is an $(d - 1)$-dimensional set (the union of a finite

number of hyperplanes) in $\mathbb{R}^d$. Similarly, let Δ_2 denote the set of vectors which are aligned with two non-parallel faces of some triangulation element. Then Δ_2 is a finite union of codimension 2 subspaces of $\mathbb{R}^d$, and in particular Δ_2 is $(d-2)$-dimensional.

Lemma 3.5.4 *Suppose q_h is transversal. If $t_{\max} < \infty$, then necessarily $\dot{q}_h(t_{\max}) \in \Delta_2$.*

Proof. If $t_{\max} < \infty$, then $t_k \to t_{\max}$ and there are infinitely many boundary crossing times t_{k_m}, for example, at which the points $q_h(t_{k_m})$ lie on the same face F of a single triangulation element. We can decompose the particle motion $q_h = q_h^{F_\perp} + q_h^F$ into scalar part $q_h^{F_\perp}$ perpendicular to F and an $(d-1)$-dimensional motion q_h^F parallel to F. Denote the two closed elements adjacent to F by F^+ and F^-.

We first show that $q_h(t_{\max})$ cannot lie in the interior of F. Suppose the contrary were true. Since $t_{k_{m+1}} - t_{k_m} \to 0$ and the boundary $\partial(F^+ \cup F^-)$ of $(F^+ \cup F^-)$ is a positive distance apart from $q_h(t_{\max})$, for sufficiently large m, q_h cannot cross this boundary in between two crossings of F. Otherwise the transversal velocity $\dot{q}_h^{F_\perp}$ would diverge. It follows that, for t large enough, $q_h(t)$ alternates between F^+ and F^-. But then the explicit form of $q_h^{F_\perp}$ as a parabola shows that the time that elapses between two crossings of F can only take two different values (depending on wether the particle moves through F^+ or F^-) and in particular does not converge to zero. This contradicts the convergence of (t_k).

So we may assume that $q_h(t_{\max}) \in \partial F$. As we have just seen that q_h cannot lie in the interior of $F^+ \cup F^-$ for all times close to $t_{\max}$, we thus get another sequence $t_{k'_m}$ such that $q_h(t_{k'_m}) \in G$, where G is another face of F^+ or F^- such that $q_h(t_{\max}) \in \partial G$.

Note that the vector $q_h(t_{k_{m+1}}) - q_h(t_{k_m})$, and thus also the difference quotient $\frac{q_h(t_{k_{m+1}}) - q_h(t_{k_m})}{t_{k_{m+1}} - t_{k_m}}$ is aligned with F. Taking the limit $m \to \infty$ we deduce that $\dot{q}_h^{F_\perp}(t_{\max}) = 0$, i.e., that also $\dot{q}_h(t_{\max})$ is aligned with this element face, since $q_h \in C^1([0, t_{\max}])$. An analogous argument with F replaced by G shows that $\dot{q}_h(t_{\max})$ is aligned with G, too. So in fact $\dot{q}_h(t_{\max})$ is aligned with $F \cap G$, and this concludes the proof. ■

Before we prove convergence of the approximating trajectories, let us note that, for transversal initial conditions, the trajectories remain transversal for a non-zero time span independent of h.

Lemma 3.5.5 *Suppose $\dot{q}_0 \notin \Delta_1$. Then there exists $T > 0$ independent of h such that $t_{\max}(h) > T$ for all h.*

Proof. Since $\ddot{q}_h(t)$ is bounded independently of h on compact intervals by Lemma 3.5.2, by choosing T small enough, we may assume that $|\dot{q}_h(t) - \dot{q}_0|$ is so small that $\dot{q}_h(t) \notin \Delta_1$ for all $t \leq \min\{T, t_{\max}(h)\}$. But then q_h is transversal on $[0, \min\{T, t_{\max}(h)\}]$, and in particular $t_{\max}(h) > T$. The claim now follows. ■

As a consequence, for transversal initial conditions this implies $\liminf_{h\to 0} t_{\max}(h) > 0$.

Theorem 3.5.1 *Fix initial conditions $(q_0, \dot{q}_0)$ such that $\dot{q}_0 \notin \Delta_1$. For all $0 < T < \liminf_{h\to 0} t_{\max}(h)$, the approximating trajectories q_h converge to the continuum trajectory q strongly in $W^{2,\infty}([0,T])$.*

Proof. By Lemma 3.5.2 $\|q_h\|_{W^{2,\infty}([0,T])}$ is bounded independently of h. Passing—if necessary—to a subsequence, we may assume that $q_h \overset{*}{\rightharpoonup} q$ in $W^{2,\infty}([0,T])$ for some $q \in W^{2,\infty}([0,T])$. But then q_h converges strongly in $W^{1,\infty}$ by the Rellich compactness theorem and since $\nabla V_h \to \nabla V$ uniformly on compacts, we may pass to the limit in (3.71) to obtain that

$$M\ddot{q} = -\nabla V(q) \qquad \text{on} \qquad [0,T] \tag{3.73}$$

Since the right hand side of (3.71) converges uniformly, also $\ddot{q}_h$ converges uniformly to $\ddot{q}$ and we obtain $q_h \to q$ strongly in $W^{2,\infty}([0,T])$.

Applying the above reasoning to an arbitrary subsequence $h_m \to 0$, we have thus proved that a further subsequence converges to a solution q of the original equation of motion. Since this solution is unique, indeed the family q_h converges to q. ■

We are now in a position to state and prove our main global convergence result. In Lemma 3.5.3 we have seen that the element boundary crossings of approximating sequences are transversal except for a $(2d-1)$-dimensional set of initial conditions. This exceptional set does indeed depend on h and, therefore, if h is viewed as a real variable we cannot expect that the set of exceptional initial values can be chosen negligibly small independently of h. However, in practice this problem does

not occur since every numerical scheme is restricted to sequences $h_m \to 0$. Under these conditions, the following theorem shows that the exceptional set is indeed negligible.

Theorem 3.5.2 *Let $h_m \to 0$. Then except for a $(2d-1)$-dimensional set of initial conditions $(q_0, \dot{q}_0)$, q_{h_m} converges to the continuum solution q strongly in $W^{2,\infty}_{\mathrm{loc}}([0,\infty))$.*

Proof. Since $\{h_m\}$ is countable, by Lemma 3.5.3 we may assume that all the element boundary crossings at times $t_1(h_m), t_2(h_m), \ldots$ are transversal. By Theorem 3.5.1 it suffices to show that $\liminf_{m\to\infty} t_{\max}(h_m) = \infty$. Suppose this were not the case, that is $\liminf_{m\to\infty} t_{\max}(h_m) = \bar{t} < \infty$. By passing to a subsequence (not relabeled) we may assume that $\bar{t} = \lim_{m\to\infty} t_{\max}(h_m)$. Now Lemma 3.5.4 implies that $\dot{q}_{h_m}(t_{\max}(h_m)) \in \Delta_2$. On the other hand, we deduce from Theorem 3.5.1 that, for all $T < \bar{t}$, $\dot{q}_{h_m}(T) \to \dot{q}(T)$. Since by Lemma 3.5.2 $\sup_{[T,t_{\max}]} \ddot{q}_{h_m}(t)$ is uniformly bounded, Δ_2 is closed and $\dot{q}$ is continuous, sending $T \to \bar{t}$ we deduce that $\dot{q}(\bar{t}) \in \Delta_2$. We conclude the proof by showing that the set of initial conditions for which $\dot{q}(t) \in \Delta_2$ at some positive time t has Hausdorff dimension $2d-1$: Let $\Phi : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ be the flow associated to the continuum equations of motion

$$\dot{q}(t) = p(t) \tag{3.74}$$

$$\dot{p}(t) = -M^{-1}\nabla V(q(t)) \tag{3.75}$$

i.e., solutions with initial conditions (q_0, p_0) are given by $(q(t), p(t)) = \Phi(t, q_0, p_0)$. The critical set of initial conditions under investigation is then given by

$$\bigcup_{t\in[0,\infty)} \Phi_t^{-1}(\mathbb{R}^d \times \Delta_2) = \Phi^{-1}(\mathbb{R}^d \times \Delta_2) \tag{3.76}$$

where $\Phi_t = \Phi(t, \cdot)$. Since $(t, q, p) \mapsto (t, \Phi_t(q, p))$ is a diffeomorphism on $\mathbb{R}^{2d+1}$ and Δ_2 is $(d-2)$-dimensional, the set $\{(t, q_0, p_0) : (t, \Phi_t(q_0, p_0)) \in \mathbb{R} \times \mathbb{R}^d \times \Delta_2\}$ is $(2d-1)$-dimensional. But $\Phi^{-1}(\mathbb{R}^d \times \Delta_2)$ is just the projection onto the first coordinate of this set. Since projections—being Lipschitz continuous—do not enlarge the dimension of a set, we have indeed $\dim \Phi^{-1}(\mathbb{R}^d \times \Delta_2) \leq 2d-1$. ∎

3.6 Numerical examples

We begin with the following quote from the survey of open problems in symplectic integration by McLachlan and Scovel [61]:

> *How efficient can symplectic integrators be, given only the Hamiltonian (Lagrangian) function?*

The answer of this question is of practical relevance either when the Lagrangian is so complicated that one does not wish to calculate its derivative by hand or when the numerical evaluation of the derivative is so computationally demanding that one does not wish to use a time integrator that relies on its calculation, e. g. multiscale modeling of materials. The force-stepping integrator defined in Section 3.2 and the continuous piecewise-linear representation of approximate energies presented in Section 3.3 result in a time integration scheme which only requires one evaluation of the potential energy in each step. Thus, force-stepping could shed some light on the question posed above.

Next, we present selected examples of application that showcase the conservation, accuracy, long-term behavior and convergence properties of force-stepping. These properties play an important role in the long-term behavior of problems with strong nonlinearities, e. g. long-term integration in celestial mechanics and computation of thermodynamics properties in molecular dynamics. For instance, experience has shown that time-step adaption provides an efficient way to cope with strong nonlinearities, but it is also observed that it tends to degrade the long-term behavior of standard time integrators [8, 20]. We illustrate the performance of force-stepping in those areas of application by means of three examples: the motion of two bodies which attract each other by Newton's law of gravitation, that is the Kepler problem, the dynamics of a frozen argon cluster, and the oblique impact of an elastic cube. In the first example, we investigate Lagrangian reductions with respect to translational and rotational symmetries. In the second example, we only investigate translational symmetry reductions by means of the general procedure presented in Section 3.4. In the third example, we do not carry out any Lagrangian reduction.

For purposes of assessing the performance of force-stepping, we draw detailed comparisons with the second-order explicit Newmark method, namely, the member of the Newmark family of time-

stepping algorithms corresponding to parameters $\beta = 0$ and $\gamma = 1/2$ (cf., e. g., [42] for a detailed account of Newmark's method). In the linear regime, explicit Newmark is second-order accurate and conditionally stable, with a critical time step equal to twice over the maximum natural frequency of the system. Explicit Newmark is identical to velocity Verlet and, for constant time step, it is also identical to central differences. It can also be shown that the Newmark solution is in one-to-one correspondence, or *shadows*, the solution of the trapezoidal-rule variational integrators (cf., e. g., [42]). Thus, explicit Newmark provides a convenient representative of a time-integrator commonly used in molecular dynamics, finite-differencing and variational integration.

Detailed analyses of the implicit members of the Newmark family of algorithms, their stability and energy preserving properties (for linear systems) were given in Belytschko and Schoeberle [4], Hughes [35] and related papers.

3.6.1 Kepler problem

3.6.1.1 Lagrangian reduction

The motion of two bodies which attract each other by Newton's inverse square law of gravitation is often called the Kepler problem. This problem is characterized by a Lagrangian $L : \mathbb{R}^6 \times \mathbb{R}^6 \to \mathbb{R}$ of the form (3.1) and is rich in symmetries and constants of motion. Among these invariants, the total linear momentum and the total angular momentum are relevant to our analysis. We first apply the ideas of Lagrangian reduction, presented in Section 3.4, to reduce the dimension of phase space to $TQ = \mathbb{R}^2 \times \mathbb{R}^2$. More precisely, we chose one of the bodies as the center of the coordinate system in order to have a two dimensional motion $q = (x, y)$. Then, the Lagrangian reduces to

$$L(x, y, \dot{x}, \dot{y}) = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2\right) + \frac{1}{\sqrt{x^2 + y^2}}. \tag{3.77}$$

and the solution is stable and periodic for initial conditions of the family

$$q_0 = (1 - e, 0) \qquad , \qquad \dot{q}_0 = \left(0, \sqrt{\tfrac{1+e}{1-e}}\right) \tag{3.78}$$

with $0 \leq e < 1$. Indeed, Kepler's first law states that planets move in elliptic orbits with the sun at one of the foci.

It is possible to further reduce phase space by first changing coordinates to a polar representation and then by eliminating the angular components, i.e., the group coordinates associated with the rotational symmetry. Then, reduction by stages gives a reduced Lagrangian $l : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ of the form (3.1),

$$l(r, \dot{r}) = \frac{1}{2}\dot{r}^2 - V'(r) \tag{3.79}$$

$$\theta(t) = \int_0^t \frac{\Theta}{r(\tau)^2} d\tau + \theta_0 \tag{3.80}$$

where the effective potential is $V'(r) = -\dfrac{1}{r} + \dfrac{\Theta^2}{2r^2}$, and the total angular momentum is $\Theta = r\dot{\theta} = x\dot{y} - y\dot{x}$. Initial conditions are

$$r_0 = 1 - e \quad , \quad \dot{r}_0 = 0 \quad , \quad \Theta = \sqrt{1 - e^2} \quad , \quad \theta_0 = 0 \tag{3.81}$$

A unique property of 2-body problems, e. g. the Kepler problem, is that reduction with respect to translational and rotational symmetries induces Lagrangians of the desirable form (3.1). Therefore, force-stepping trajectories can be readily obtained either from (3.77) or from (3.79). The latter are linear and angular momenta preserving and the former are linear momentum preserving. This is in sharp contrast to N-body problems with $N \geq 3$, which only allow for an effective reduction with respect to the translational symmetry, as described in Section 3.4.

We first consider the Kepler problem described by the Lagrangian (3.77) and initial conditions (3.78) with $e = 0.85$. The size of the underlying simplicial triangulation employed in the force-stepping calculations is $h_{x,y} = 0.022$. The time step employed in the time-stepping calculations is $\Delta t = 0.0125$, which corresponds to the average time step resulting from the force-stepping calculations. The total duration of the analysis is 64π.

The constant time-step explicit Newmark method is symplectic-momentum preserving and, therefore, does not conserve energy. As it is often the case with symplectic-momentum preserving meth-

ods, it nevertheless has good energy-conservation properties for sufficiently small time steps. However, due to the conditional stability of the method the time-step is constrained by the maximum natural frequency of the system. The one-step explicit algorithm is

$$q_{k+1} = q_k + \Delta t \dot{q}_k - \frac{\Delta t^2}{2} M^{-1} \nabla V(q_k) \tag{3.82a}$$

$$\dot{q}_{k+1} = \dot{q}_k - \frac{\Delta t}{2} M^{-1} \left[\nabla V(q_k) + \nabla V(q_{k+1}) \right] \tag{3.82b}$$

By way of contrast, LaBudde and Greenspan [50] presented a second-order accurate and symmetric method which is energy-momentum preserving and, therefore, does not conserve the symplectic structure. This time integrator was designed as a modification of the midpoint rule method considering a potential energy of the form $V(q) = U(|q|)$, e. g. the Kepler problem satisfies this restrictive condition. The one-step implicit algorithm is

$$q_{k+1} = q_k + \frac{\Delta t}{2} \left(\dot{q}_k + \dot{q}_{k+1} \right) \tag{3.83a}$$

$$\dot{q}_{k+1} = \dot{q}_k - \Delta t \frac{U(|q_{k+1}|) - U(|q_k|)}{|q_{k+1}|^2 - |q_k|^2} M^{-1} \left(q_{k+1} + q_k \right) \tag{3.83b}$$

Figure 3.4 shows the trajectories provided by force-stepping, explicit Newmark, and the energy-momentum preserving time integrator. Although a precession effect is characteristic of these numerical trajectories, we note that orbits are stable and elliptic over long times. The good long-time behavior of force-stepping and explicit Newmark is due to their exact conservation of the symplectic structure. In contrast, the good long-time behavior of the energy-momentum method is understood as a consequence of its time-reversibility more than of its exact conservation properties [30]. Conservation of total energy and/or angular momentum of these methods are verified in Figure 3.5. It is also worth noting that, at least for this dynamic system of interest, there is no apparent *drift* in the angular momentum of force-stepping trajectories. This may be understood as a consequence of the near-momentum conservation properties of the method (Theorem 3.4.2).

Figure 3.6 shows a histogram of the time steps selected by force-stepping. The broad range of

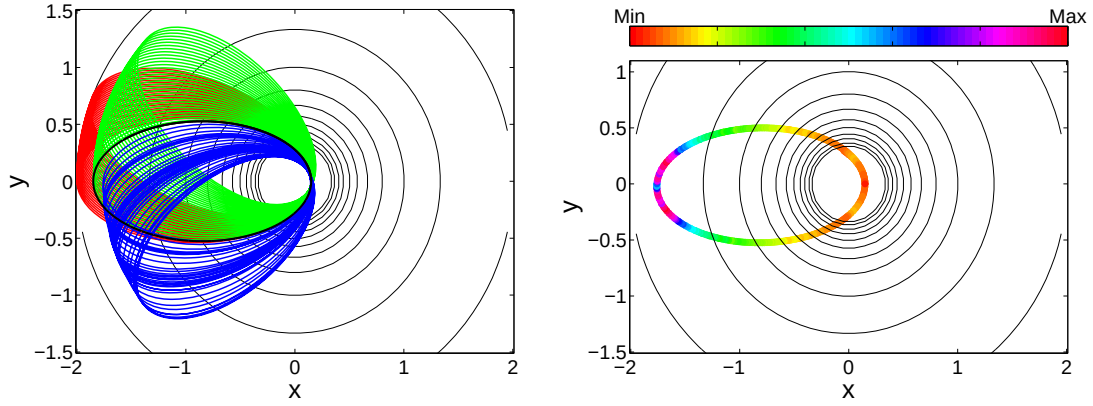


Figure 3.4: Kepler problem described by $L(x, y, \dot{x}, \dot{y})$ with $e = 0.85$ and $\Delta t = 0.0125$. Left: Force-stepping (blue line), Newmark (red line), energy-momentum (green line) and exact (black line) trajectories in configuration space. Right: Spatial distribution of time steps selected by force-stepping for $h_{x,y} = 0.022$.

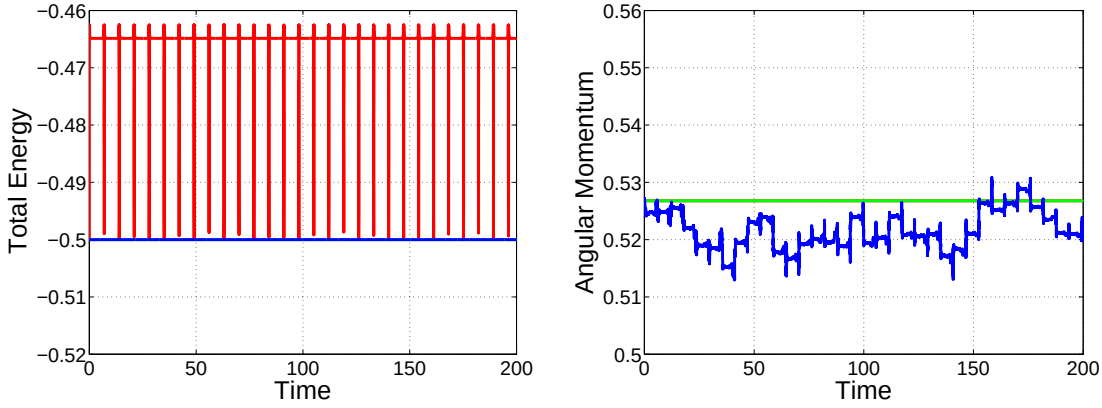


Figure 3.5: Kepler problem described by $L(x, y, \dot{x}, \dot{y})$ with $e = 0.85$. Blue line: Force-stepping solution with $h_{x,y} = 0.022$. Red line: Newmark solution with $\Delta t = 0.0125$. Green line: Energy-momentum solution with $\Delta t = 0.0125$.

time steps is noteworthy, as is their distribution in configuration space shown on the right side of Figure 3.4. In regions where the velocity is high (low), the resulting time steps are small (large). Thus, theses figures also illustrate the automatic time-selection property of force-stepping. This property may in turn be regarded as the means by which force-stepping achieves symplecticity, exact conservation of energy, and approximate conservation of momentum maps. It also bears emphasis that, unlike variable time-step variational integrators designed to conserve energy [40, 51], force-stepping always selects a valid time step and is therefore free of solvability concerns.

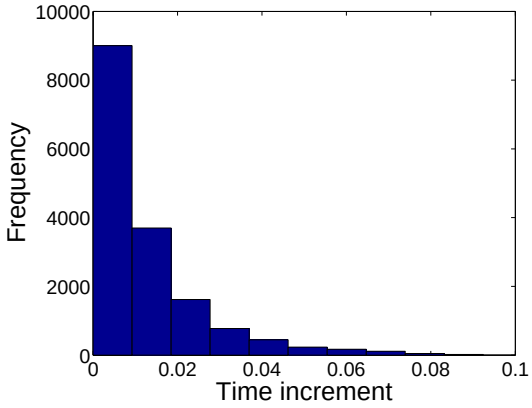


Figure 3.6: Kepler problem described by $L(x, y, \dot{x}, \dot{y})$ with $e = 0.85$. Histogram of time steps selected by force-stepping for $h_{x,y} = 0.022$.

We now consider the Kepler problem described by the reduced Lagrangian (3.79) and initial conditions (3.81) with $e = 0.85$. The size of the underlying simplicial triangulation employed in the force-stepping calculations is $h_r = 0.0067$—we note that the simplicial triangulation reduces to a one-dimensional grid. The time step employed in the time-stepping calculations is $\Delta t = 0.0125$, which corresponds to the average time step resulting from the force-stepping calculations as well as to the time step employed in the above calculations. The total duration of the analysis is 64π.

Figure 3.7 shows the reconstructed trajectories provided by force-stepping and explicit Newmark. The original dynamics is reconstructed from the reduced dynamics by computing $\theta(t)$ from (3.80). The angular displacement can be exactly obtained from force-stepping trajectories and therefore the total angular momentum map is conserved along the reconstructed trajectories. Conservation of total energy and angular momentum of the reconstructed trajectories are verified in Figure 3.8. As

demonstrated in Section 3.4, force-stepping is a symplectic-energy time-reversible integrator which, by recourse to Lagrangian reduction, conserves exactly the total linear and angular momentum maps of the original Lagrangian.

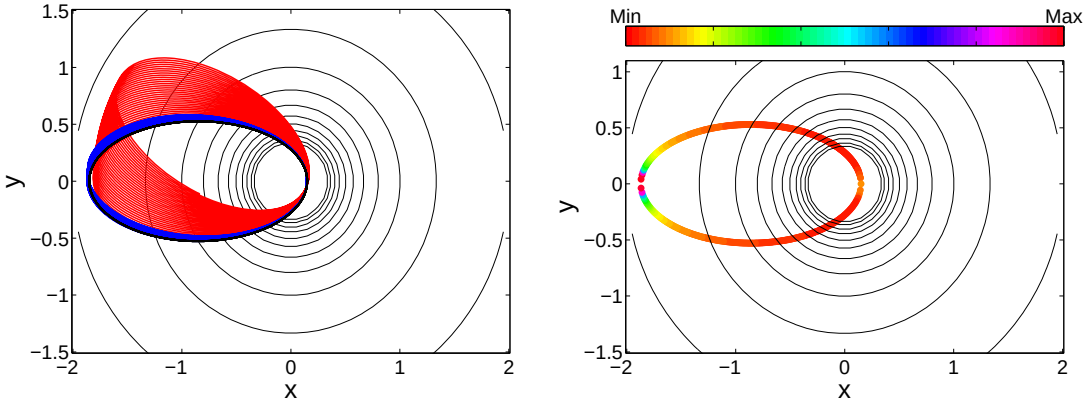


Figure 3.7: Kepler problem described by $l(r, \dot{r})$ with $e = 0.85$ and $\Delta t = 0.0125$. Left: Force-stepping (blue line), Newmark (red line) and exact (black line) reconstructed trajectories in configuration space. Right: Spatial distribution of time steps selected by force-stepping for $h_r = 0.0067$.

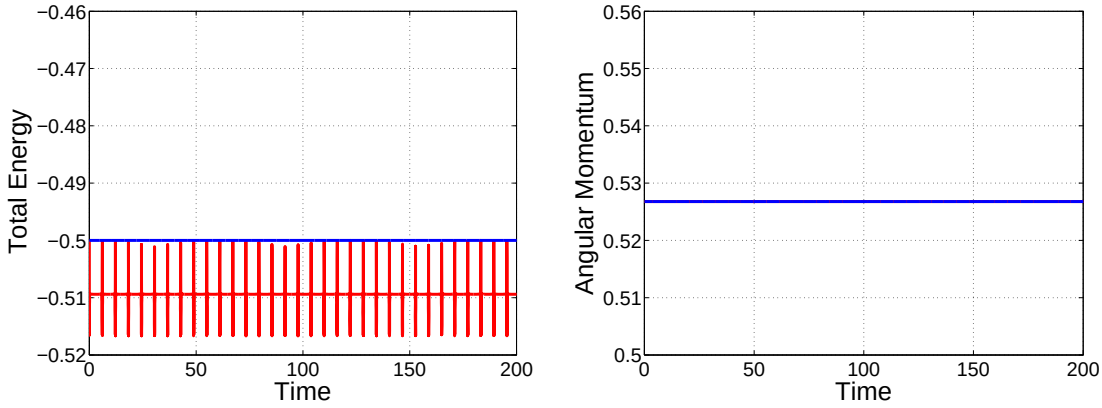


Figure 3.8: Kepler problem described by $l(r, \dot{r})$ with $e = 0.85$. Blue line: Force-stepping solution with $h_r = 0.0067$. Red line: Newmark solution with $\Delta t = 0.0125$.

It is suggestive to note that the automatic time-selection property of force-stepping leads to a sharper distribution of time steps in the reduced configuration space than in the original configuration space (cf. the right side of Figure 3.4 and the right side of Figure 3.7). The direct relationship between group symmetries of the system and optimal time-step adaption is noteworthy.

3.6.1.2 Strong nonlinearities

McLachlan and Scovel [61] recognize the Kepler problem with $e \to 1$ as a strong nonlinear problem and suggest taking this case as a good test of a variable time step method. We then consider the Kepler problem described by the Lagrangian (3.77) and initial conditions (3.78) with $e = 0.99$. Thus, the potential energy gradient goes from $\nabla V(q_0) = (10^4, 0)$ at the initial configuration to $\nabla V(q_a) \simeq (1, 0)$ at the apoapsis—farthest point from the focus. The size of the underlying simplicial triangulation employed in the force-stepping calculations is $h_{x,y} = 0.000247$. The time step employed in the time-stepping calculations is $\Delta t = 0.000175$, which corresponds to the average time step resulting from the force-stepping calculations. The total duration of the analysis is 16π.

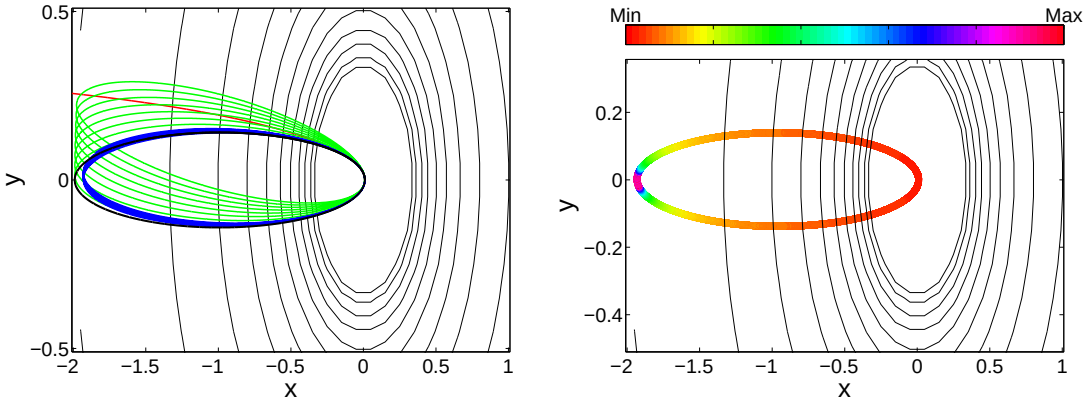


Figure 3.9: Kepler problem described by $L(x, y, \dot{x}, \dot{y})$ with $e = 0.99$ and $\Delta t = 0.000175$. Left: Force-stepping (blue line), Newmark (red line), energy-momentum (green line) and exact (black line) trajectories in configuration space. Right: Spatial distribution of time steps selected by force-stepping for $h_{x,y} = 0.000247$.

Figure 3.9 shows force-stepping, explicit Newmark, and energy-momentum preserving trajectories in configuration space. We first note that the orbits of both force-stepping and energy-momentum methods are stable and elliptic over long times, while explicit Newmark does not even exhibit the qualitative periodic behavior of the solution. We also note that a larger precession effect of the energy-momentum preserving integrator is in clear detriment of its pointwise accuracy. The remarkable behavior of force-stepping relays on its automatic time-step selection required for achieving symplecticity, exact conservation of energy, and approximate conservation of momentum maps. Indeed, the direct relationship between the spatial distribution of time steps and the nonlinearity of

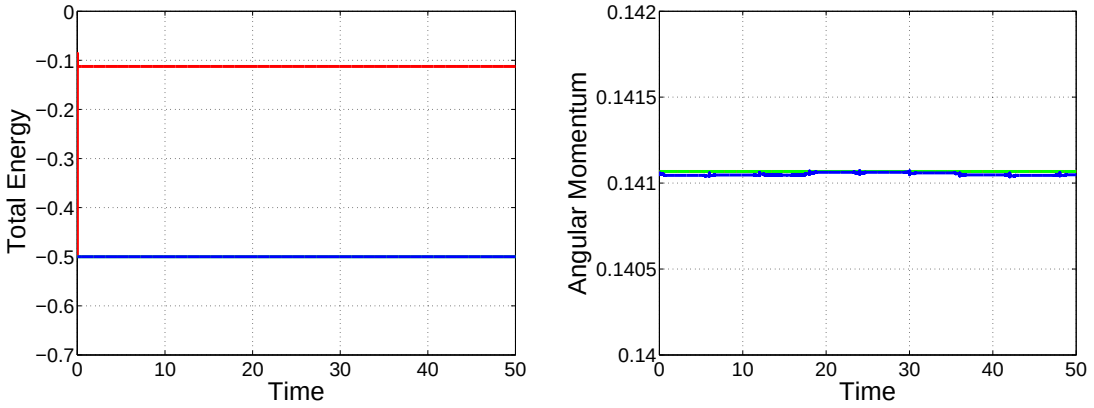


Figure 3.10: Kepler problem described by $L(x, y, \dot{x}, \dot{y})$ with $e = 0.99$ and $\Delta t = 0.000175$. Force-stepping (blue line), Newmark (red line) and energy-momentum (green line) solutions.

the problem—visualized by the iso-curves of potential energy—is illustrated on the right side of Figure 3.9. In contrast, the good behavior of the energy-momentum method may relay on its implicit algorithm specially designed for central field forces, i.e. $V(q) = U(|q|)$, which is significantly more computationally expensive than force-stepping. Finally, conservation of total energy and angular momentum are verified in Figure 3.10.

3.6.1.3 Convergence analysis

We motivate this section with the following problem posed by McLachlan and Scovel [61]:

> *Develop variable time step symplectic integrators so that they are competitive for pointwise accuracy with standard methods for the Kepler problem, while retaining the good long-time behavior of constant time step symplectic methods.*

To this end, we first recall those characteristic parameters that describe the exact solution of the Kepler problem, i.e., the semi-major axis $a = L_0^2/(1 - e^2) = 1$, the semi-minor axis $b = (1 - e^2)^{1/2}$ and the period $\omega = 2\pi$. In the spirit of phase error analysis, these are the *statistical* quantities relevant to a global accuracy analysis rather than individual trajectories over intermediate to long time scales.

Figure 3.11 shows convergence results for $a_h \to a = 1$ and $\omega_h \to \omega = 2\pi$, i.e., the *spectral convergence*. Figure 3.12 shows convergence results for $|q_h - q|_\infty \to 0$, i.e., *pointwise convergence*. We note that the force-stepping is one order of magnitude more accurate than explicit Newmark

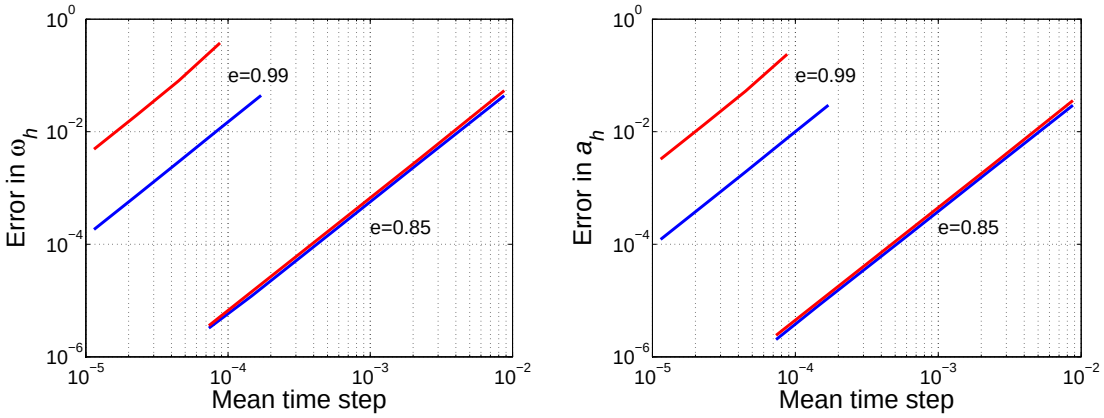


Figure 3.11: Kepler problem described by $L(x, y, \dot{x}, \dot{y})$ with $e = 0.85$ and $e = 0.99$. Blue line: Force-stepping solution. Red line: Newmark solution.

when the problem has strong nonlinearities ($e = 0.99$). We also note that both methods have the same convergence rate and exhibit nearly identical accuracy characteristics when the problem has moderate nonlinearities ($e = 0.85$). These results not only indicate that standard time integrators, e. g. explicit Newmark, are particularly not efficient nor effective in solving problems with strong nonlinearities during large periods of time, but they also demonstrate that force-stepping significantly outperforms classical methods under these conditions.

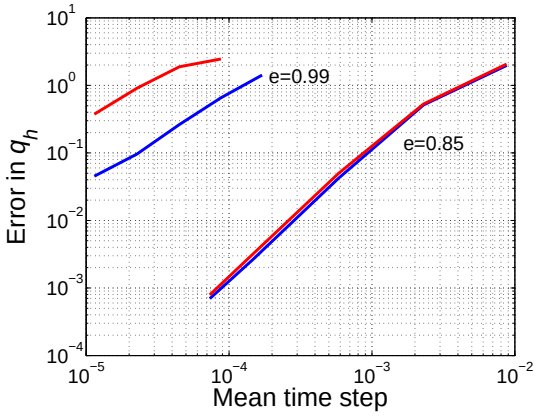


Figure 3.12: Kepler problem described by $L(x, y, \dot{x}, \dot{y})$ with $e = 0.85$ and $e = 0.99$. Blue line: Force-stepping solution. Red line: Newmark solution.

3.6.2 Frozen argon cluster

Molecular dynamics falls squarely within the framework considered in this chapter. Many applications in materials science, such as the calculation of free energies, require the integration of the system over long periods of time. In these applications, it is essential that the time integrator have good long-time behavior, such as conferred by symplecticity and exact conservation properties.

The velocity Verlet scheme is perhaps the most widely used time-integration scheme in molecular dynamics. As already noted, velocity Verlet is identical to explicit Newmark and, for constant time steps, it is symplectic-momentum preserving with good energy-conservation properties for sufficiently small time steps. However, due to the conditional stability of the method the time-step is constrained by the period of thermal vibrations of the atoms, which renders calculations of equilibrium thermodynamic properties exceedingly costly. The development of integration schemes that alleviate or entirely eliminate the time-step restrictions of explicit integration in molecular dynamics applications is the subject of ongoing research (cf., e. g., [36, 38]).

We proceed to illustrate the performance of force-stepping in molecular dynamics applications by analyzing the dynamics of a simple argon cluster. Specifically, we consider the numerical experiment proposed by Biesiadecki and Skeel [5]. The experiment concerns the two-dimensional simulation of a seven-atom argon cluster, six atoms of which are arranged symmetrically around the remaining central atom, Figure 2.5. The atoms interact via the pairwise Lennard-Jones potential

$$\phi(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{3.84}$$

where r is the distance between two atomic centers, $\varepsilon/k_B = 119.8$ K and $\sigma = 0.341$ nm are material constants for Argon, and $k_B = 1.380658 \cdot 10^{-23}$ J/K is the Boltzmann's constant. In addition, the mass of an argon atom is $m = 66.34 \cdot 10^{-27}$ kg. The initial positions of the atoms are slightly perturbed about the configuration that minimizes the potential energy of the cluster. The initial velocities are chosen such that the total linear momentum is zero and the center of mass of the cluster remains fixed. The corresponding total energy of the cluster is $E/\varepsilon = -10.51928$. Table 2.1

summarizes the initial conditions for the simulation.

Two different discretizations of configuration space are employed in the force-stepping calculations: $h_1 = 0.020$ nm and $h_2 = 0.005$ nm. The time steps employed in the velocity Verlet calculations are $\Delta t_1 = 3.12$ fsec and $\Delta t_2 = 0.80$ fsec. These time steps correspond to the average time steps resulting from the respective force-stepping calculations. The total duration of the analysis is 10 nsec.

The first row of Figs. 3.13 and 3.14 show the evolution in time of the constants of motion of the system: i) total energy, ii) total linear momentum, and iii) total angular momentum. As expected, force-stepping is energy and linear momentum preserving regardless the discretization of configuration space employed. It is worth noting that, while not being exactly angular momentum conserving, force-stepping exhibits good angular momentum behavior over long times.

In order to assess the long-time behavior of force-stepping, we present on the second row of Figs. 3.13 and 3.14 the qualitative behavior of the trajectories: iv) the evolution in time of the numerical temperature of the cluster, v) a histogram of time steps selected by the time integrator, and vi) the trajectories of the argon atoms for a time window of $[9.95, 10]$ nsec (the configuration at $t = 0$ is represented by the dashed line hexagon). The numerical temperature of the cluster is computed, under the assumption of ergodicity, by

$$T(t) = \frac{1}{tNk_B} \int_0^t \frac{1}{2} m \|\dot{q}(\tau)\|^2 d\tau \tag{3.85}$$

where $N = 7$ is the number of particles of the cluster. We note that force-stepping does not suffer from numerical dissipation and the system quickly reaches a thermodynamic temperature—the thermodynamic temperature is defined in the limit of $t \to \infty$. We also note that trajectories of the seven argon atoms, each represented by a different color, remain stable over long periods of time.

Finally, the convergence statement of Theorem 3.5.2 is illustrated by Figs. 3.15 and 3.16. We first recall that, as demonstrated in Section 3.5, the approximate space of trajectories is $X = W^{2,\infty}_{\text{loc}}(0, T)$.

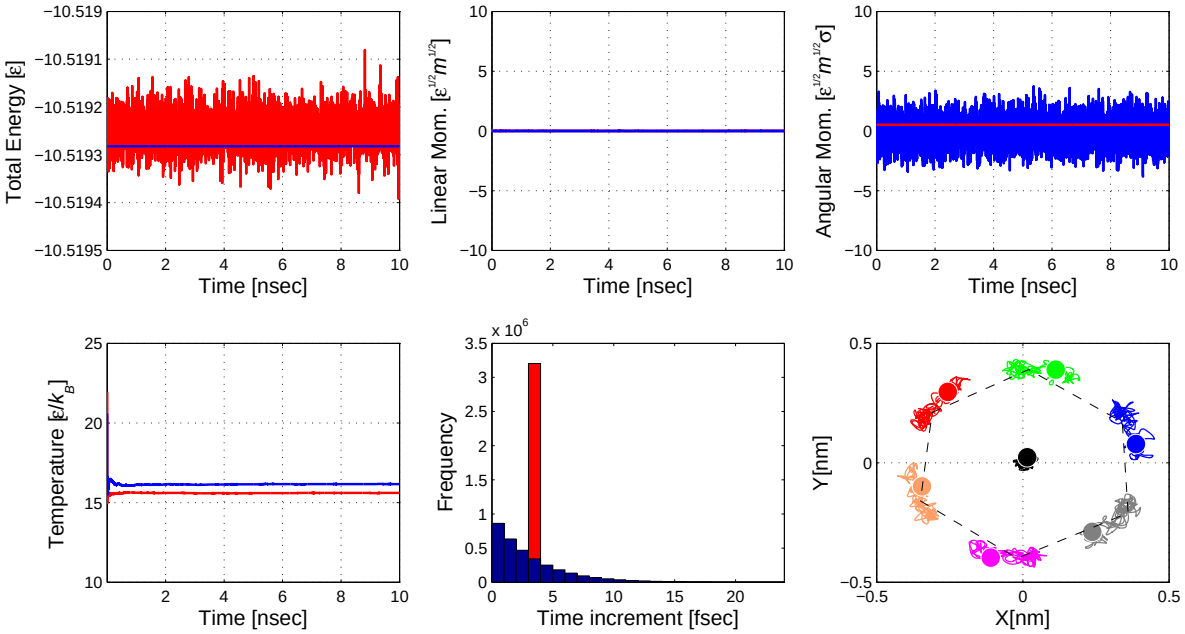


Figure 3.13: Frozen argon cluster. Blue line: Force-stepping solution with $h_1 = 0.02$ nm. Red line: Newmark solution with $\Delta t_1 = 3.12$ fsec.

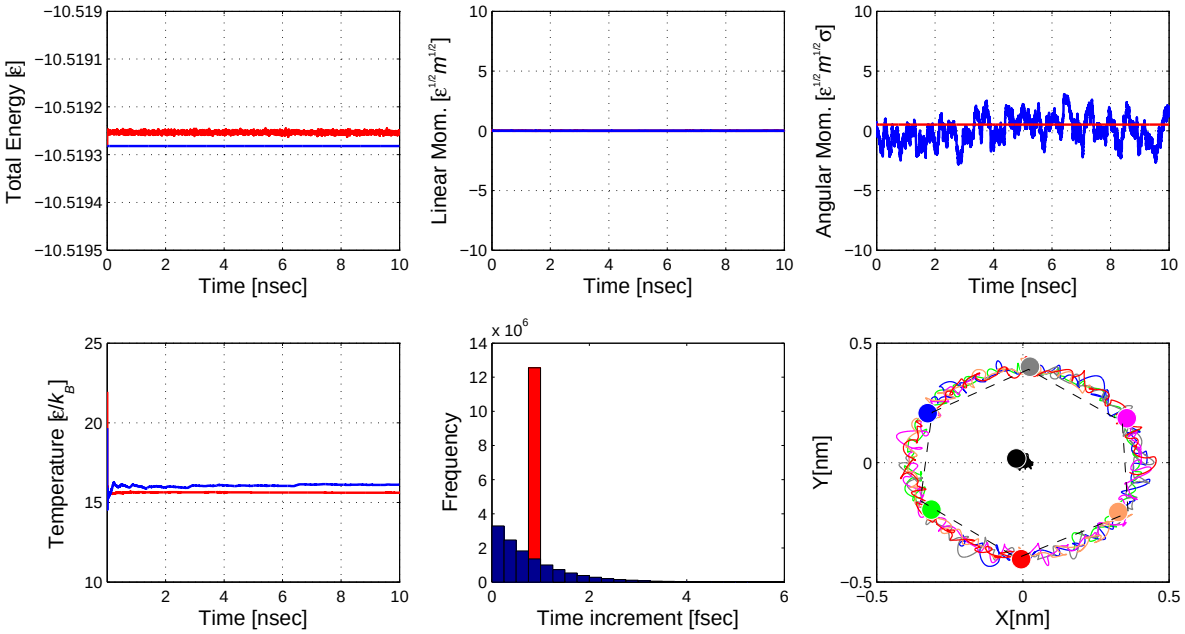


Figure 3.14: Frozen argon cluster. Blue line: Force-stepping solution with $h_2 = 0.005$ nm. Red line: Newmark solution with $\Delta t_2 = 0.80$ fsec.

Then, distances are measured with respect to the $W^{2,\infty}$-norm

$$\|q\|_{W^{2,\infty}(0,T)} = \|q(t)\|_{L^\infty} + \|\dot{q}(t)\|_{L^\infty} + \|\ddot{q}(t)\|_{L^\infty} \tag{3.86}$$

and a relative $W^{2,\infty}$-error is defined by

$$W^{2,\infty}\text{-error} = \frac{\big|\|q_h\|_{W^{2,\infty}(0,T)} - \|q\|_{W^{2,\infty}(0,T)}\big|}{\|q\|_{W^{2,\infty}(0,T)}} \tag{3.87}$$

where $\|q\|_{W^{2,\infty}(0,T)}$ is estimated from numerical results. The relative $W^{2,\infty}$-error of force-stepping with and without parity-invariance is shown on the right side of Figs. 3.15 and 3.16, respectively, as a function of the simplicial grid size h, for $T = 0.1$ nsec. The slope of the convergence plot is also directly related to the rate of convergence in grid size, that is $W^{2,\infty}$-error $= O(h^r)$. This gives an estimated rate of convergence in the $W^{2,\infty}$-norm of $r \simeq 1$. Furthermore, it is interesting to observe on the left side of Figs. 3.15 and 3.16 that the average and the maximum time steps selected by force-stepping are $O(h)$. Additionally, the minimum time step selected by the method is $O(h^{7/4})$ when an approximate potential V_h with parity-invariance is employed, and $O(h^2)$ when the approximate potential breaks the parity symmetry of the exact potential V. It is also interesting to note that force-stepping with parity-invariance exhibits a region of asymptotic convergence larger than force-stepping without parity-invariance. Therefore, the sequence q_h is indeed convergent and transversal, as expected from the analysis in Section 3.5, and force-stepping with parity-invariance exhibits better convergence properties.

3.6.3 Finite-element model with contact: oblique impact of neo-Hookean cube

Next we consider finite-dimensional Lagrangian systems obtained by a finite-element discretization of the action of a nonlinear elastic solid (cf., e. g., [51] for details of finite-element approximation in elastodynamics). For these applications, the generalized coordinates q of the system are the coordinates of the nodes in the deformed configuration of the solid. The example that we present

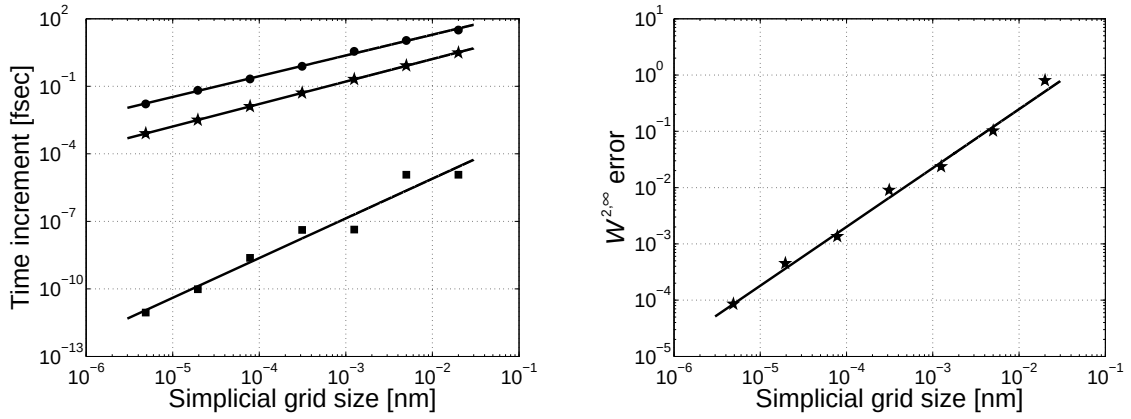


Figure 3.15: Convergence analysis of the frozen argon cluster using force-stepping with parity-invariance. Right: Convergence is observed in the $W^{2,\infty}$-norm with estimated convergence rate of $r \simeq 1$. Left: The average ($\star$) and the maximum ($\bullet$) time steps selected by force-stepping are $O(h)$, whereas the minimum ($\blacksquare$) time step is $O(h^{7/4})$.

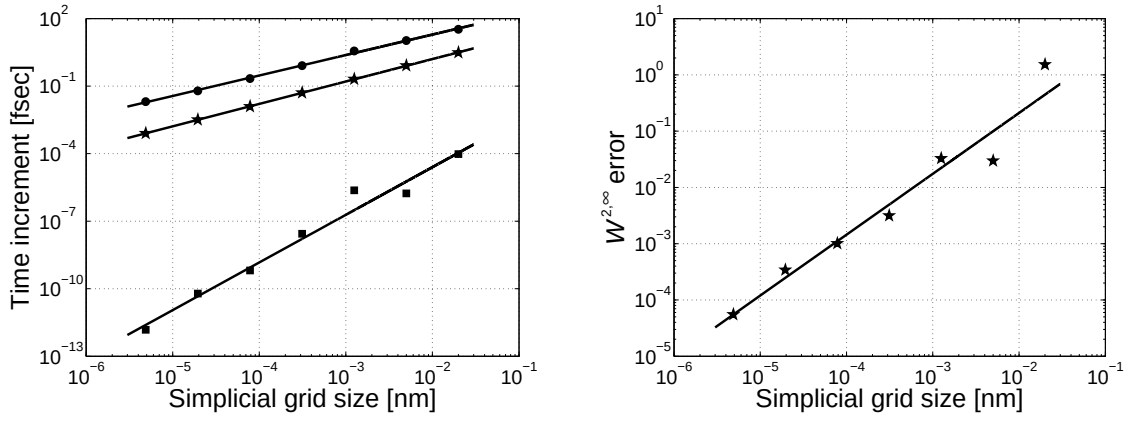


Figure 3.16: Convergence analysis of the frozen argon cluster using force-stepping without parity-invariance. Right: Convergence is observed in the $W^{2,\infty}$-norm with estimated convergence rate of $r \simeq 1$. Left: The average ($\star$) and the maximum ($\bullet$) time steps selected by force-stepping are $O(h)$, whereas the minimum ($\blacksquare$) time step is $O(h^2)$.

to illustrate the performance of force-stepping in that area of application is the oblique impact of a neo-Hookean cube in a rigid wall. To this end, we assume a strain-energy density of the form

$$W(F) = \frac{\lambda_0}{2}(\log J)^2 - \mu_0 \log J + \frac{\mu_0}{2}\operatorname{tr}\left(F^T F\right) \tag{3.88}$$

which describes a neo-Hookean solid extended to the compressible range. In this expression, λ_0 and μ_0 are Lamé constants, and J is the Jacobian of the deformation gradient F.

In problems involving contact we additionally consider the kinematic restrictions imposed by the impenetrability constraint. We recall that the admissible configuration set $\mathcal{C}$ of a deformable body is the set of deformation mappings which are globally one-to-one. In so-called *barrier* methods, the interpenetration constraint may be accounted for by adding the indicator function $I_{\mathcal{C}}(q)$ of the admissible set $\mathcal{C}$ to the energy of the solid. We recall that the indicator function of a set $\mathcal{C}$ is the extended-valued function

$$I_{\mathcal{C}}(q) = \begin{cases} 0 & \text{if } q \in \mathcal{C} \\ \infty & \text{otherwise} \end{cases} \tag{3.89}$$

Often in calculations, the indicator function $I_{\mathcal{C}}$ is replaced by a penalty approximation $I_{\mathcal{C},\epsilon} \geq 0$ parameterized by a small parameter $\epsilon > 0$ and such that $I_{\mathcal{C},\epsilon} = 0$ over $\mathcal{C}$. In this approach, as $\epsilon \to 0$, $I_{\mathcal{C},\epsilon} \to I_{\mathcal{C}}$ pointwise and interpenetration is increasingly penalized. A convenient choice of penalty energy function for contact with a rigid hyperplane is of the form

$$I_{\mathcal{C},\epsilon}(q) = \frac{1}{2\epsilon}\sum_{\alpha \in I} g_\alpha(q) \tag{3.90}$$

where the index set I ranges over all boundary nodes and

$$g_\alpha(q) = \begin{cases} 0, & \text{if } (q-O)\cdot n \geq 0, \\ \|(q-O)\cdot n\|^2, & \text{otherwise,} \end{cases} \tag{3.91}$$

where O and n are the hyperplane reference point and outer-pointing normal. We note that the admissible set $I_{\mathcal{C},\epsilon}$ is not invariant under the action of translations and rotations. It therefore

follows that the constrained Lagrangian does not retain its momentum preserving properties and the Lagrangian reduction described in Section 3.4 cannot be carried out.

The application of force-stepping to dynamic contact problems is straightforward. The case in which the constraints are represented by means of a penalty energy function $I_{\mathcal{C},\epsilon}$ falls right within the general framework and requires no special considerations, that is $V(q) = W(F(q)) + I_{\mathcal{C},\epsilon}(q)$. In addition, collision is automatically captured by the intrinsic time adaption of force-stepping (cf., e. g., [10, 80], for a detailed discussion of time-step selection considerations in contact problems).

Our example concerns the oblique impact of an elastic aluminum cube of size equal to 0.1 m. The mesh comprises 192 4-node tetrahedral isoparametric elements and 71 nodes. The cube is a compressible neo-Hookean solid characterized by a strain-energy density of the form (3.88). The values of the material constants in (3.88) are: $\lambda_0 = 60.5$ GPa, $\mu_0 = 26$ GPa, and $\rho = 2700$ kg/m^3. The cube is imparted an initial velocity $v_0 = (1, 1, 1)$ km/s. The rigid wall is described by a hyperplane with $O = (0.101, 0.101, 0.101)$ m and $n = (0, -1, -1)$. A sequence of snapshots of the force-stepping trajectory corresponding to $h = 5 \cdot 10^{-4}$ m are shown on Figs. 3.17 and 3.18.

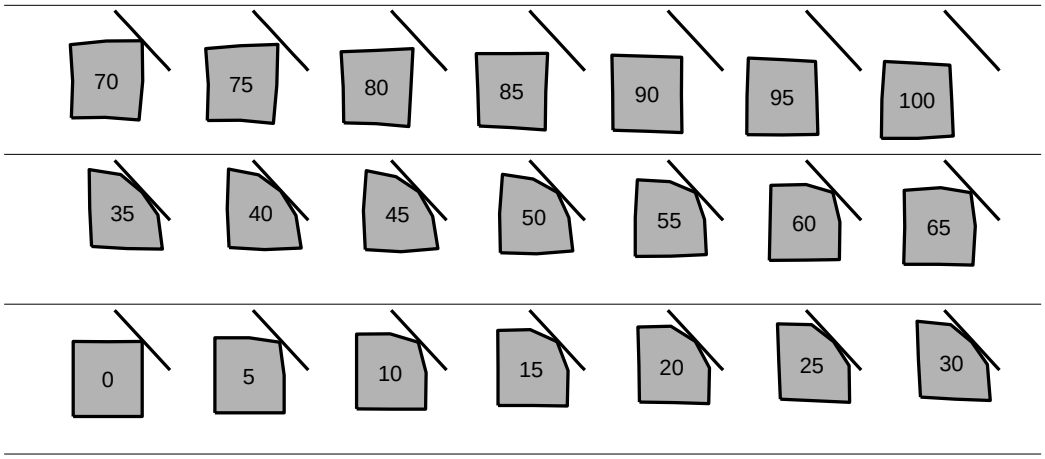


Figure 3.17: Oblique impact of neo-Hookean cube. Sequence of snapshots of the force-stepping trajectory with $h = 5 \cdot 10^{-4}$ m displaying the yz-midplane of the cube. Numbers indicate times in μs and oblique lines represent the rigid wall.

Figure 3.19 compares the time histories of total energy, potential energy and kinetic energy attendant to the force-stepping trajectory for $h = 5 \cdot 10^{-4}$ m and Newmark's trajectory for $\Delta t =$ 0.2 nsec, the latter chosen to be in the order of the average time step resulting from the force-

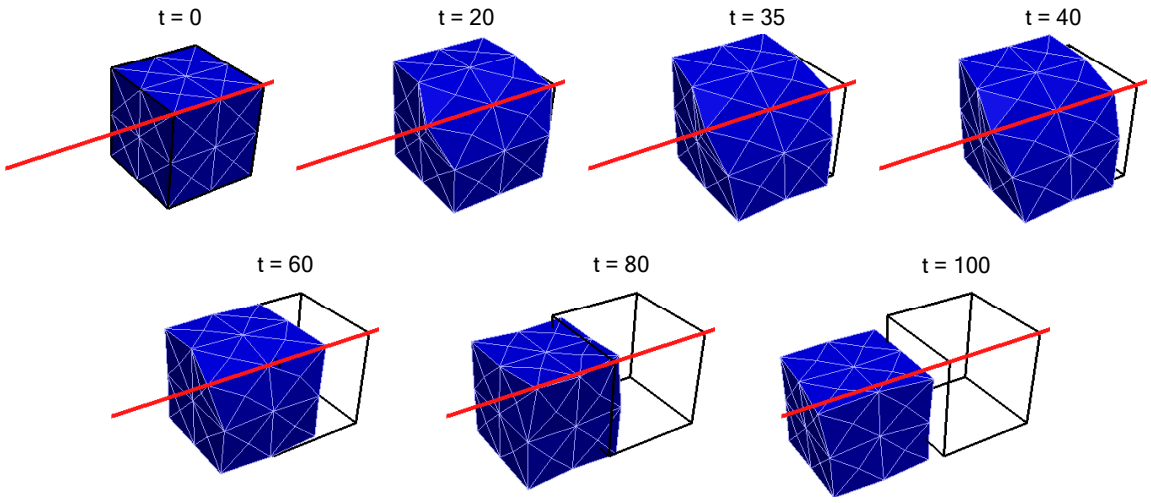


Figure 3.18: Oblique impact of neo-Hookean cube. Sequence of snapshots of the force-stepping trajectory with $h = 5 \cdot 10^{-4}$ m. Numbers indicate times in μs and red lines provide a reference for the first points of impact, that is the $(x, 0.101, 0.101)$-line.

stepping calculations. As expected, the kinetic energy of the system is partly converted to potential energy during the approach part of the collision sequence, and vice versa during the release part. Also characteristically, force-stepping is observed to conserve energy exactly through the collision, whereas the Newmark energy remains within tight bounds due to the small time step and the appropriate penalty energy, i.e. (3.90) and (3.91) with $\epsilon = 10^{-14}$. Other notable features of the calculations are the ability of force-stepping to detect the time of collision and to automatically modulate the time-step so as to resolve the fine structure of the intricate interactions that occur through the collision.

Figure 3.20 shows the time history of the linear and angular momenta of the cube. Because the cube impacts with a rigid wall, the total linear and angular momenta of the system are not conserved through the collision but evolve in a characteristic way with time. We note from Figure 3.20 that force-stepping trajectory does indeed follow closely Newmark trajectory, even though force-stepping does not conserve linear and angular momenta. We also note that the transfer of linear momentum from the rigid wall to the cube is in the direction normal to the hyperplane and the motion in the orthogonal direction is unperturbed. Similarly, the oblique impact introduces angular momentum into the system. The good momentum behavior of force-stepping is noteworthy and

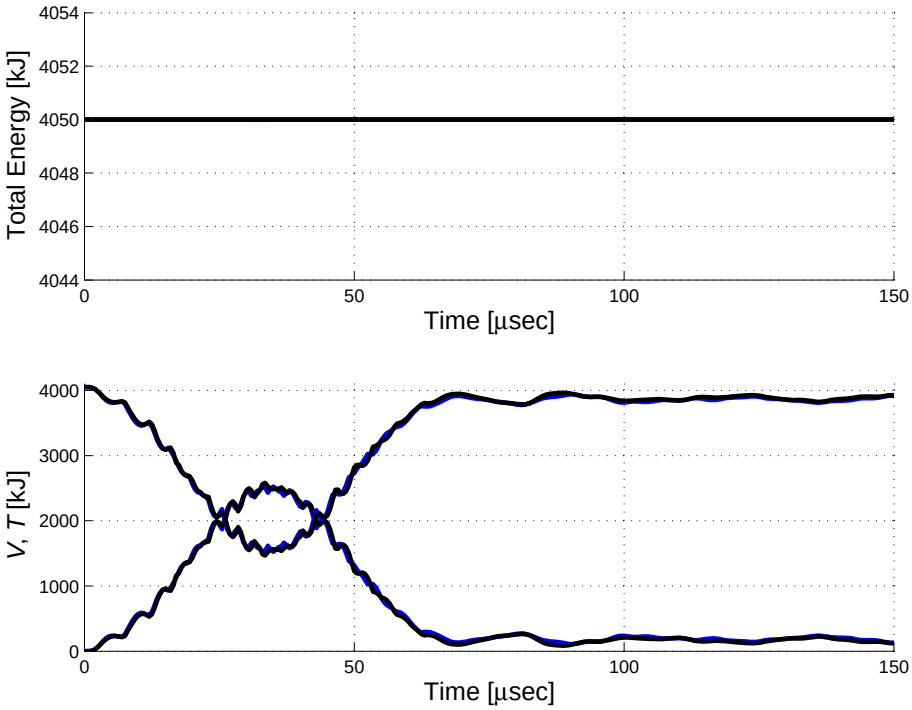


Figure 3.19: Oblique impact of neo-Hookean cube. Blue line: Force-stepping solution with $h = 5 \cdot 10^{-4}$ m. Black line: Newmark solution with $\Delta t = 0.2$ nsec.

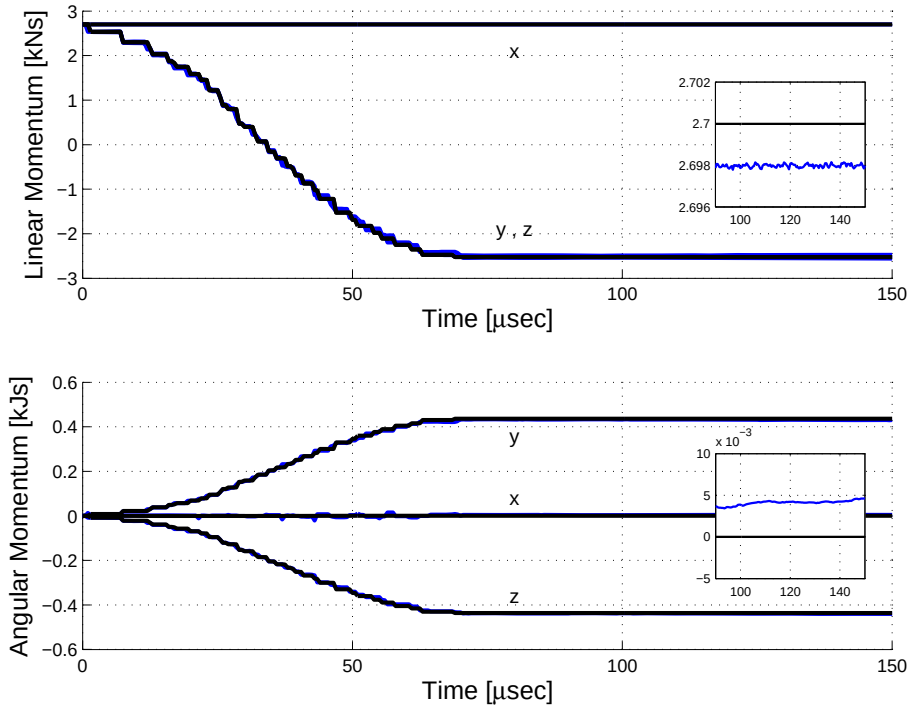


Figure 3.20: Oblique impact of neo-Hookean cube. Blue line: Force-stepping solution with $h = 5 \cdot 10^{-4}$ m. Black line: Newmark solution with $\Delta t = 0.2$ nsec. The insets show in detail the x-component of momentum maps.

may be understood as a consequence of the near-momentum conservation properties of the method (Theorem 3.4.2).

Finally, we investigate the relationship between the average time step and the number of degrees of freedom of the system $d = 3N$. To this end we consider three different meshes with increasing level of refinement, i.e. meshes I, II and II with 213, 375 and 1287 degrees of freedom, respectively. Figure 3.21 shows the average time step selected by force-stepping is $O(h/N^2)$, where h is the simplicial grid size. We note that these results are consistent with and extend the results obtained in Section 3.6.2, i.e. a dependence of the form $O(h)$.

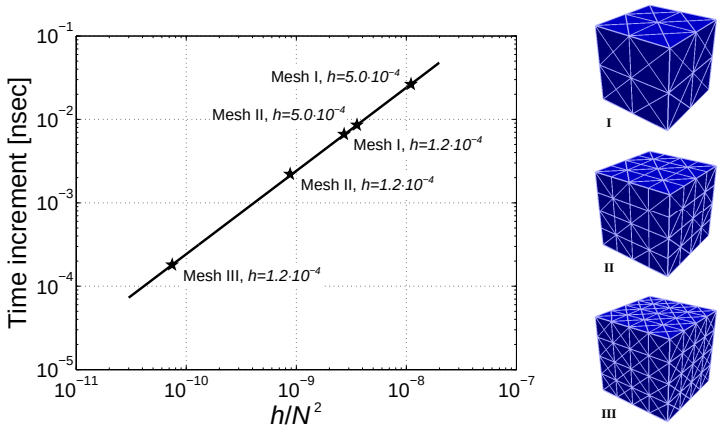


Figure 3.21: Oblique impact of neo-Hookean cube. The average time step selected by force-stepping is $O(h/N^2)$, where h is the simplicial grid size and $3N$ is the number of degrees of freedom of the system. Three different meshes are considered and depicted on the right hand side.

3.7 Summary and discussion

We have formulated a class of integration schemes for Lagrangian mechanics, which we refer to as *force-stepping integrators*, that are symplectic-energy and time-reversible with automatic selection of the time step size. The scheme also conserves approximately all the momentum maps associated with the symmetries of the system. Exact conservation of momentum maps may be achieved by recourse to Lagrangian reduction. The general strategy leading to the formulation of the force-stepping scheme, and its forerunner, the energy-stepping scheme [24], may be viewed as the reverse of backward-error analysis. Thus, whereas backward-error analysis seeks to identify a nearby Lagrangian system

that is solved exactly by the solutions generated by a numerical integrator, the approach followed here is to directly replace the system by a nearby one that can be solved exactly. The force-stepping scheme is obtained by replacing the original—or reduced—potential energy by a piecewise affine approximation over a simplicial grid, or regular triangulation. By taking triangulations of diminishing size, an approximating sequence of energies is generated. The trajectories of the resulting approximate Lagrangians can be characterized explicitly and consist of piecewise parabolic motion, or *free fall*. We have shown that the force-stepping trajectories are symplectic, energy preserving, approximately conserve momentum maps of the original system and, except for a negligible small set of initial conditions, converge to trajectories of the original system when the size of the simplicial grid is decreased to zero. Selected numerical tests, including the Kepler problem, the dynamics of a frozen argon cluster, and the oblique impact of an elastic cube, demonstrate the excellent long-term behavior of force-stepping, its automatic time-step selection property, and the ease with which it deals with constraints, including contact problems.

We have also described a unique, systematic and efficient representation of the piecewise-linear approximation of the potential based on a regular triangulation of $\mathbb{R}^d$ into proper simplices. The interpolation of the approximate potential can be restricted to one simplex at a time, and the *replacement rule* provides an efficient scheme to construct adjacent simplices as the numerical solution crosses simplicial boundaries. In particular, we have developed an algorithm for updating all simplex-related matrices required to compute force-stepping trajectories that has complexity $O(d^2)$.

We close by pointing out some limitations of our analysis and possible avenues for extensions of the approach.

Firstly, our analysis of near conservation of momentum maps holds for given time intervals $[0, T]$ with $T < \infty$, that is the momentum map of force-stepping trajectories converges to the original one in $W^{1,\infty}_{\text{loc}}([0, \infty))$. It is possible that such analysis can be extended to $W^{1,\infty}([0, \infty))$ if conditions under which approximating trajectory errors average out over long times are understood. Our experience with selected numerical tests suggests that the nearly-conserved momentum maps remain within tight bounds for long periods of time. However, a rigorous analysis of this property is beyond the

scope of this work.

Secondly, automatic time-step selection is an attractive feature of the force-stepping scheme. We have shown a linear relationship between the simplicial grid size and the average time step. In practice, this scaling could be improved by using triangulations that require fewer simplices. The review of Brandts *et al.* [7] on nonobtuse simplicial partitions, the work of Bliss and Su [6] on lower bounds for simplicial covers and triangulations of cubes, and other related work suggest possible directions in that regard.

Thirdly, piecewise polynomial interpolation of the potential, including piecewise constant interpolation (*energy-stepping* [24]) and piecewise linear interpolation (*force-stepping*) do not exhaust the class of approximations that generate exactly solvable Lagrangians. A case in point concerns the use discontinuous piecewise-linear approximations of the potential energy over a grid of polytopes. However, criteria for constructing the piecewise-linear patches and convergence properties of the resulting time-integrator are not well-understood at present. The systematic investigation of approximation schemes of the type proposed here, the elucidation of their properties and the determination of the best types of approximating Lagrangians in each area of application, are worthwhile directions of future research.

Chapter 4

Asynchronous energy-stepping integrators

We present a symplectic-energy-momentum time integrator for Lagrangian mechanics, referred to as *asynchronous energy-stepping* integrator, that has the ability to automatically and asynchronously modulate the time step in different regions of the domain, if any. In order to achieve these properties we introduce a localization of the configuration space, or domain partition, that allows for an additive decomposition of the *global* Lagrangian into *local* Lagrangians amenable to energy-stepping approximations. Thus, we replace the original *localized* potential energies by stepwise or *terraced* approximations at steps of uniform heights. The trajectories of the resulting Lagrangians can be characterized explicitly and consist of intervals of rectilinear motion that span consecutive level contours of the localized potential energies. We show that the asynchronous energy-stepping trajectories are symplectic, exactly conserve all momentum maps of the original local systems and account for the exact momentum fluxes across subdomain boundaries. These properties, the excellent long-term behavior of asynchronous energy-stepping and its automatic and asynchronous time-step selection property are born out by selected examples of application, including the head-on collision of soft elastic spherical balls and the propagation of solitary waves in a chain of elastic beads.

4.1 Introduction

In recent papers [24, 25] and Chapters 2-3, the authors have proposed a method of approximation for Lagrangian mechanics consisting of replacing the Lagrangian $L(q, \dot{q})$ of the system by a sequence of approximate Lagrangians $L_h(q, \dot{q})$ that can be *solved exactly*. The approximate solutions $q_h(t)$ are then the *exact* trajectories of $L_h(q, \dot{q})$. In this manner, the approximate solutions are themselves trajectories of a Lagrangian system and, therefore, have many of the properties of such trajectories such as symplecticity and exact conservation of the momentum maps associated with the symmetry groups of $L(q, \dot{q})$ preserved by $L_h(q, \dot{q})$. For example, in [24], this program has been carried out in full for Lagrangians of the type: $L(q, \dot{q}) = K(\dot{q}) - V(q)$, with $K(\dot{q})$ a quadratic form, and for approximate Lagrangians of the type $L_h(q, \dot{q}) = K(\dot{q}) - V_h(q)$, with $V_h(q)$ a terraced piecewise constant approximation to $V(q)$. This type of approximation does indeed result in exactly solvable Lagrangians which have all the symmetries of the original system. The approximate trajectories then exactly conserve all the momentum maps of the system, whether explicitly known (e. g., total energy, linear and angular momenta) or not. It bears emphasis that, in contrast to variational integrators (cf., e. g., [52, 53, 58]), the momenta that are conserved are the exact momenta of the system and not some time discretization thereof. The approximate trajectories are piecewise rectilinear, with the intervals of rectilinear motion spanning consecutive level contours of the potential, hence the name *energy-stepping* integrators. The durations of these intervals may be regarded as time steps, whose determination is part of the solution process. In this manner, the approach overcomes an intrinsic limitation of fixed time-step integrators, which cannot simultaneously conserve energy, the symplectic structure and other conserved quantities, such as linear and angular momenta [19]. Additionally, under mild restrictions on the potential, the approximate trajectories are found to indeed converge to exact trajectories of the system, subject to technical transversality constraints.

Many applications in material science and structural mechanics are characterized by dynamical systems which exhibit a wide range of strain rates and time scales that coexist in time and space (e. g., many-body contact dynamics). Moreover, in most cases the interesting dynamics takes place in a small region of the domain and this region of interest drifts in space as time progresses (e. g.,

propagation of shock waves, solitary waves, structural and material instabilities, or fracture dynamics). In these applications, an effective and efficient time integrator must have good long-term behavior—such as conferred by symplecticity and exact conservation properties—and the ability to automatically and asynchronously modulate the time step in different regions or subdomains. Some of these properties are shared by subcycling methods [34, 63], group implicit algorithms [67], asynchronous variational integrators [51], energy-stepping integrators [24], and related works. Energy-stepping integrators, in particular, are convergent time-integration schemes with exact conservation and automatic time-step selection properties. In this chapter, we endow energy-stepping with asynchronous time-step selection properties, hence the term *asynchronous energy-stepping* integrators. We achieve these desirable properties by a localization of the configuration space that allows for an additive decomposition of the original *global* Lagrangian into *local* Lagrangians amenable to energy-stepping approximations (for example, in the context of finite-elements models, the localization is often the result of partitioning the mesh into subdomains). By this simple procedure, the approximate trajectories are symplectic and exactly preserve all conservation laws of the original local systems. Furthermore, the approximate trajectories are piecewise rectilinear, with intervals of rectilinear motion that span consecutive level contours of the *localized* potential energies. Thus, the durations of these intervals differ in each subdomain and may be regarded as *local* time steps, whose asynchronic determination is part of the solution process.

The chapter is organized as follows. The asynchronous energy-stepping time-integration scheme is defined in Section 4.2. The global and local conservation properties of asynchronous energy-stepping are presented in Section 4.3. In Section 4.4 we discuss some aspects of the implementation and computational efficiency of the method. Finally, in Section 4.5 we present selected examples of application that illustrate the efficiency, convergence, accuracy and conservation properties of asynchronous energy-stepping, including the head-on collision of soft elastic spherical balls and the propagation of solitary waves in a chain of elastic beads. Finally, a summary and concluding remarks are collected in Section 4.6.

4.2 Asynchronous energy-stepping integrators

For definiteness, we specifically consider dynamical systems characterized by Lagrangians $L : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ of the form

$$L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V(q) \tag{4.1}$$

where M is the mass matrix and V is the potential energy. Lagrangians of this form arise in a number of areas of application including structural dynamics and molecular dynamics. Let us now define n_{LV} *local* variables $q^{(j)} \in \mathbb{R}^{m_j < d}$ as the localization of the configuration q, that is

$$q^{(j)} = Lo^{(j)} q \tag{4.2}$$

where $Lo^{(j)} \in \mathbb{R}^{m_j \times d}$ is the j-th localization matrix. It is clear that $\{q^{(1)} \dots q^{(n_{LV})}\}$ carries the same information as q, since the former is obtained from the latter as simple process of localization. In the context of finite-elements models, the localization is often the result of partitioning the mesh, i.e., each extended variable corresponds to a subdomain in the partition. In the context of coupled systems with an interface (e. g., many-body contact dynamics) domain decomposition methods apply in a natural way, since subdomains are already defined. In molecular dynamics, the localization is often achieved by decomposing the potential energy into contributions with different characteristic time and length scales. Let us also assume that the dynamical systems in consideration allow for an additive decomposition of the Lagrangian into local Lagrangians $L^{(j)}(q^{(j)}, \dot{q}^{(j)})$ of the form (4.1), that is

$$L(q, \dot{q}) = \sum_{j=1}^{n_{LV}} L^{(j)}(q^{(j)}, \dot{q}^{(j)}) = \sum_{j=1}^{n_{LV}} \left\{ \frac{1}{2}(\dot{q}^{(j)})^T M^{(j)} \dot{q}^{(j)} - V^{(j)}(q^{(j)}) \right\} \tag{4.3}$$

Then, the localized mass matrix $M^{(j)}$ and potential energy $V^{(j)}$ naturally verify the following relations

$$M = \sum_{j=1}^{n_{LV}} (Lo^{(j)})^T M^{(j)} Lo^{(j)} \tag{4.4}$$

$$V(q) = \sum_{j=1}^{n_{LV}} V^{(j)}(q^{(j)}) \tag{4.5}$$

Following the method proposed by the authors in [24,25], the trajectories of a Lagrangian system can be approximated by replacing $L(q, \dot{q})$ by an approximating Lagrangian $L_h(q, \dot{q})$ that can be solved exactly. A particular type of approximating Lagrangian is

$$L_h(q^{(j)}, \dot{q}^{(j)}) = \sum_{j=1}^{n_{LV}} \frac{1}{2} (\dot{q}^{(j)})^T M^{(j)} \dot{q}^{(j)} - \sum_{j=1}^{n_{LV}} V_h^{(j)}(q^{(j)}) \tag{4.6}$$

obtained by introducing an approximation of the potential energy. In this work, we specifically investigate piecewise constant approximations of the localized potential energies. Thus, we replace the original localized potential energies $V^{(j)}$ by stepwise or *terraced* approximations $V_h^{(j)}$ at steps of uniform height $h^{(j)}$, namely,

$$V_h^{(j)}(q^{(j)}) = h^{(j)} \lfloor (h^{(j)})^{-1} V^{(j)}(q^{(j)}) \rfloor \tag{4.7}$$

where $\lfloor \cdot \rfloor$ is the floor function, i.e., $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}$. By taking steps of diminishing height, an approximating sequence of potential energies and Lagrangians is generated in this manner. The chief characteristics of the new systems thus obtained are that they can be solved exactly, as demonstrated in Section 4.2.1, and that they have the same symmetries as the original system, as shown in Section 4.3. It bears emphasis that asynchronous energy-stepping integrators result from the piecewise constant approximation of the localized potential energy $V^{(j)}(q^{(j)})$, whereas energy-stepping integrators result from the piecewise constant approximation of the *global* potential energy of the system $V(q)$ [24].

4.2.1 Computation of the exact trajectories of the approximating Lagrangian

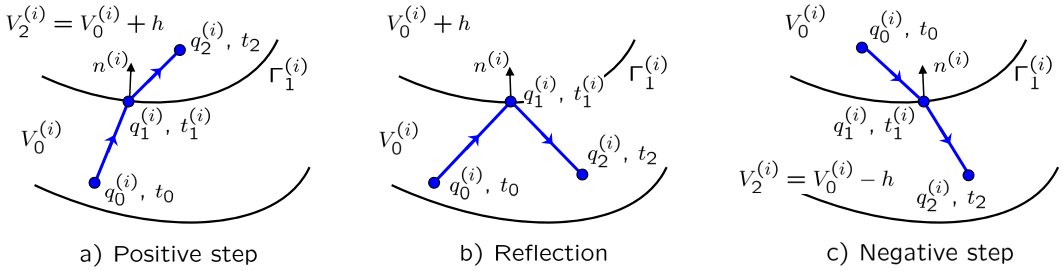


Figure 4.1: Trajectory of a system whose localized potential energy is approximated as piecewise constant.

We now proceed to compute the exact trajectories of the approximate Lagrangian $L_h(q^{(j)}, \dot{q}^{(j)})$ resulting from the piecewise constant approximation (4.7) of the potential energy. Suppose that the system is in configuration $\{q_0^{(1)} \dots q_0^{(j)} \dots q_0^{(n_{LV})}\}$ at time t_0 and in configuration $\{q_2^{(1)} \dots q_2^{(j)} \dots q_2^{(n_{LV})}\}$ at time t_2 and that during the time interval $[t_0, t_2]$ the system intersects one single jump surface $\Gamma_1^{(i)}$ separating two regions of constant energies $V_0^{(i)}$ and $V_2^{(i)}$, Figure 4.1. By the construction of $V_h^{(i)}$, $\Gamma_1^{(i)}$ is the level surface $V^{(i)} = V_2^{(i)}$ for an uphill step $V_2^{(i)} = V_0^{(i)} + h^{(i)}$, or the level surface $V^{(i)} = V_0^{(i)}$ for a downhill step, $V_2^{(i)} = V_0^{(i)} - h^{(i)}$. For simplicity, we shall further assume that $V^{(i)}$ is differentiable and that all energy-level crossings are transversal, i.e.,

$$n^{(i)}(q_1^{(i)}) \cdot \dot{q}_1^{(i)-} \neq 0 \tag{4.8}$$

where $\dot{q}_1^{(i)-} = \dot{q}^{(i)}(t_1^{(i)-})$ and $n^{(i)}(q_1^{(i)})$ is a vector normal to $\Gamma_1^{(i)}$ pointing in the direction of advance.

Under the preceding assumptions, the action integral over the time interval $[t_0, t_2]$ follows as

$$I_h = \int_{t_0}^{t_2} L_h(q^{(j)}, \dot{q}^{(j)})\ dt = \int_{t_0}^{t_1^{(i)}} L_h(q^{(j)}, \dot{q}^{(j)})\ dt + \int_{t_1^{(i)}}^{t_2} L_h(q^{(j)}, \dot{q}^{(j)})\ dt \tag{4.9}$$

where $t_1^{(i)}$ is the time at which the trajectory intersects $\Gamma_1^{(i)}$. In regions where $V_h^{(j)}(q)$ is constant the trajectory $q^{(j)}(t)$ is linear in time. Therefore, the action of the system can be computed exactly

and reduces to

$$I_h = \left(t_1^{(i)} - t_0\right) \sum_{j=1}^{n_{LV}} \left\{ \frac{1}{2} \left(\frac{q_1^{(j)} - q_0^{(j)}}{t_1^{(i)} - t_0} \right)^T M^{(j)} \left(\frac{q_1^{(j)} - q_0^{(j)}}{t_1^{(i)} - t_0} \right) - V_0^{(j)} \right\} + \left(t_2 - t_1^{(i)}\right) \sum_{j=1}^{n_{LV}} \left\{ \frac{1}{2} \left(\frac{q_2^{(j)} - q_1^{(j)}}{t_2 - t_1^{(i)}} \right)^T M^{(j)} \left(\frac{q_2^{(j)} - q_1^{(j)}}{t_2 - t_1^{(i)}} \right) - V_2^{(j)} \right\} \quad (4.10)$$

where $q_1^{(i)} = q^{(i)}(t_1^{(i)})$ is constrained to be on the jump surface $\Gamma_1^{(i)}$. Assuming differentiability of $\Gamma_1^{(i)}$, stationarity of the action with respect to $(t_1^{(i)}, q_1)$ additionally gives the *global* energy conservation equation

$$\sum_{j=1}^{n_{LV}} \left\{ \left(\frac{q_1^{(j)} - q_0^{(j)}}{t_1^{(i)} - t_0} \right)^T M^{(j)} \left(\frac{q_1^{(j)} - q_0^{(j)}}{t_1^{(i)} - t_0} \right) + 2V_0^{(j)} - \left(\frac{q_2^{(j)} - q_1^{(j)}}{t_2 - t_1^{(i)}} \right)^T M^{(j)} \left(\frac{q_2^{(j)} - q_1^{(j)}}{t_2 - t_1^{(i)}} \right) - 2V_2^{(j)} \right\} = 0 \quad (4.11)$$

and the *global* linear momentum balance equation

$$\sum_{j=1}^{n_{LV}} \left\{ (Lo^{(j)})^T M^{(j)} \frac{q_1^{(j)} - q_0^{(j)}}{t_1^{(i)} - t_0} - (Lo^{(j)})^T M^{(j)} \frac{q_2^{(j)} - q_1^{(j)}}{t_2 - t_1^{(i)}} \right\} + \lambda_1 (Lo^{(i)})^T n^{(i)}(q_1^{(i)}) = 0 \quad (4.12)$$

where λ_1 is a Lagrange multiplier.

In order to make a more direct connection with time-integration schemes we reformulate the problem slightly by assuming that t_0, q_0—the latter on a jump surface $\Gamma_0^{(i)}$ except, possibly, at the initial time—and the initial velocity

$$\dot{q}_0^+ = \dot{q}\left(t_0^+\right) = \frac{q_1 - q_0}{t_1^{(i)} - t_0} \quad (4.13)$$

are known. Let $t_1^{(i)}$ and q_1 be the time and point at which the trajectory intersects the next jump surface $\Gamma_1^{(i)}$, i.e., $V^{(i)}(Lo^{(i)} q_1) = h^{(i)}\mathbb{Z}$. We then seek to determine

$$\dot{q}_1^+ = \dot{q}(t_1^+) \quad (4.14)$$

A reformulation of equations (4.11) and (4.12) in terms of $\dot{q}_1^+$ gives

$$(\dot{q}_1^+)^T M \dot{q}_1^+ = (\dot{q}_1^-)^T M \dot{q}_1^- - 2\Delta V \tag{4.15}$$

$$\dot{q}_1^+ = \dot{q}_1^- + \lambda_1 M^{-1} n_1 \tag{4.16}$$

where $\dot{q}_1^- = \dot{q}_0^+$, and the energy jump ΔV and the normal vector n_1 are

$$\Delta V = V_h^{(i)}(q^{(i)}(t_1^+)) - V_h^{(i)}(q^{(i)}(t_1^-)) \tag{4.17}$$

$$n_1 = (Lo^{(i)})^T n^{(i)}(q_1^{(i)}) \tag{4.18}$$

Next we proceed to parse through the various cases that arise in the solution of (4.15) and (4.16).

4.2.1.1 Diffraction by downhill energy step

Suppose that $\Delta V = -h^{(i)}$, i.e., the system decreases its energy as the trajectory crosses $\Gamma_1^{(i)}$. Then (4.15) becomes

$$\lambda_1^2 (n_1^{(i)})^T M^{-1(i)} n_1^{(i)} + 2\lambda_1 \dot{q}_1^{(i)-} \cdot n_1^{(i)} - 2h^{(i)} = 0 \tag{4.19}$$

where the inverse mass matrix is localized as follows

$$M^{-1(i)} = Lo^{(i)} M^{-1} (Lo^{(i)})^T \tag{4.20}$$

We note that the localized inverse mass matrix is not equal to the inverse of the localized mass matrix, that is $M^{-1(i)} \neq (M^{(i)})^{-1}$. Then, the system (4.16 - 4.19) has the real solution

$$\dot{q}_1^{(j)+} = \dot{q}_1^{(j)-} + \lambda_1 Lo^{(j)\cap(i)} M^{-1(i)} n_1^{(i)} \tag{4.21}$$

where

$$\lambda_1 = \frac{-\dot{q}_1^{(i)-} \cdot n_1^{(i)} - \sqrt{\left(\dot{q}_1^{(i)-} \cdot n_1^{(i)}\right)^2 + 2h^{(i)} \left((n_1^{(i)})^T M^{-1(i)} n_1^{(i)}\right)}}{(n_1^{(i)})^T M^{-1(i)} n_1^{(i)}} \tag{4.22}$$

and $Lo^{(j)\cap(i)} = Lo^{(j)}(Lo^{(i)})^{-1} \in \mathbb{R}^{m_j \times m_i}$ stands for the boundary mapping of local variables $q^{(i)}$ and $q^{(j)}$ (e. g., $Lo^{(j)\cap(i)}$ reduces to the identity mapping if $j = i$, and to the null mapping if $q^{(j)}$ and $q^{(i)}$ are disjoint subsets of q). This solution represents the diffraction, or change of direction, of the trajectory by a downhill energy step.

4.2.1.2 Diffraction by uphill energy step

Suppose now that $\Delta V = h^{(i)}$, i.e., the system increases its energy as the trajectory crosses $\Gamma_1^{(i)}$. Then (4.15) becomes

$$\lambda_1^2 (n_1^{(i)})^T M^{-1(i)} n_1^{(i)} + 2\lambda_1 \dot{q}_1^{(i)-} \cdot n_1^{(i)} + 2h^{(i)} = 0 \tag{4.23}$$

Suppose, in addition, that

$$\left(\dot{q}_1^{(i)-} \cdot n_1^{(i)}\right)^2 > 2h^{(i)} \left((n_1^{(i)})^T M^{-1(i)} n_1^{(i)}\right) \tag{4.24}$$

Then, the system (4.16 - 4.23) again has a real solution, namely,

$$\dot{q}_1^{(j)+} = \dot{q}_1^{(j)-} + \lambda_1 Lo^{(j)\cap(i)} M^{-1(i)} n_1^{(i)} \tag{4.25}$$

where

$$\lambda_1 = \frac{-\dot{q}_1^{(i)-} \cdot n_1^{(i)} + \sqrt{\left(\dot{q}_1^{(i)-} \cdot n_1^{(i)}\right)^2 - 2h^{(i)} \left((n_1^{(i)})^T M^{-1(i)} n_1^{(i)}\right)}}{(n_1^{(i)})^T M^{-1(i)} n_1^{(i)}} \tag{4.26}$$

This solution again represents the diffraction of the trajectory by an uphill energy step when the system has sufficient initial energy to overcome the energy barrier.

4.2.1.3 Reflection by uphill energy step

Suppose now that $\Delta V = h^{(i)}$, i.e., the system increases its energy as the trajectory crosses $\Gamma_1^{(i)}$, but, contrary to the preceding case,

$$\left(\dot{q}_1^{(i)-} \cdot n_1^{(i)}\right)^2 < 2h^{(i)} \left((n_1^{(i)})^T M^{-1(i)} n_1^{(i)}\right) \tag{4.27}$$

Then, the system (4.16 - 4.23) has no real solutions, showing that the system does not have sufficient initial kinetic energy to overcome the energy barrier. Instead, the trajectory remains within the same energy level and equation (4.15) becomes

$$(\dot{q}_1^{(i)+})^T M \dot{q}_1^{(i)+} = (\dot{q}_1^{(i)-})^T M \dot{q}_1^{(i)-} \tag{4.28}$$

Then, the system (4.16 - 4.28) is solved by

$$\dot{q}_1^{(j)+} = \dot{q}_1^{(j)-} + \lambda_1 Lo^{(j)\cap(i)} M^{-1(i)} n_1^{(i)} \tag{4.29}$$

where

$$\lambda_1 = 2\frac{\dot{q}_1^{(i)-} \cdot n_1^{(i)}}{(n_1^{(i)})^T M^{-1(i)} n_1^{(i)}} \tag{4.30}$$

This solution represents the reflection of the trajectory by an uphill energy step when the system does not have sufficient initial energy to overcome the energy barrier.

4.2.2 Summary of the asynchronous energy-stepping scheme

We close this section by summarizing the relations obtained in the foregoing and defining the *asynchronous energy-stepping* approximation scheme resulting from a piecewise-constant approximation of each localized potential energy.

Definition 4.2.1 (Asynchronous energy-stepping) Suppose $\left(t_k^{(i)}, q_k^{(i)}, \dot{q}_k^{(i)+}\right)$ and a piecewise-constant approximation of the potential energy $V_h^{(i)}$ are given. Let $t_{k+1}^{(i)}$ and $q_{k+1}^{(i)}$ be the time and points of exit of the rectilinear trajectory $q_k^{(i)} + (t^{(i)} - t_k^{(i)})\dot{q}_k^{(i)+}$ from the set $\{V^{(i)} = h^{(i)}\mathbb{Z}\}$. Let $\Delta V^{(i)}$ be the energy jump at $q_{k+1}^{(i)}$ in the direction of advance. The updated velocity is, then,

$$\dot{q}_{k+1}^{(j)+} = \dot{q}_k^{(j)+} + \lambda_{k+1} Lo^{(j)\cap(i)} M^{-1(i)} n_{k+1}^{(i)} \tag{4.31}$$

where

$$\lambda_{k+1} = \begin{cases} -2\dfrac{\dot{q}_k^{(i)+} \cdot n_{k+1}^{(i)}}{(n_{k+1}^{(i)})^T M^{-1(i)} n_{k+1}^{(i)}}, \\ \qquad\qquad \text{if } \left(\dot{q}_k^{(i)+} \cdot n_{k+1}^{(i)}\right)^2 < 2\Delta V^{(i)} \left((n_{k+1}^{(i)})^T M^{-1(i)} n_{k+1}^{(i)}\right), \\ \dfrac{-\dot{q}_k^{(i)+} \cdot n_{k+1}^{(i)} + \text{sign}\left(\Delta V^{(i)}\right) \sqrt{\left(\dot{q}_k^{(i)+} \cdot n_{k+1}^{(i)}\right)^2 - 2\Delta V^{(i)} \left((n_{k+1}^{(i)})^T M^{-1(i)} n_{k+1}^{(i)}\right)}}{(n_{k+1}^{(i)})^T M^{-1(i)} n_{k+1}^{(i)}}, \\ \qquad\qquad \text{otherwise,} \end{cases} \tag{4.32}$$

These relations define an asynchronous discrete propagator

$$\Phi_h : \left(t_k^{(i)}, q_k^{(i)}, \dot{q}_k^{(i)+}\right) \mapsto \left(t_{k+1}^{(i)}, q_{k+1}^{(i)}, \dot{q}_{k+1}^{(i)+}\right) \tag{4.33}$$

that can be iterated to generate a discrete trajectory. We provide in Section 4.4 details of its implementation and computational efficiency.

4.3 Conservation properties

We first recall that energy-stepping [24] exactly conserves all the momentum maps and the symplectic structure of the original Lagrangian system because the stepwise approximation of the potential energy preserves all the symmetries of the system and the discrete trajectories are exact trajectories of a Lagrangian system. Likewise, as we show in this section, asynchronous energy-stepping exactly conserves all the momentum maps and the symplectic structure of the original local Lagrangian system.

In order to facilitate the analysis we recast Definition 4.2.1 into its sequential or synchronous form. Then, the potential energy is additively decomposed (4.5) and stepwise approximated (4.7) but a single sequence of times is employed (i.e., $\{t_k\} = \bigcup_j \{t_k^{(j)}\}$). Indeed, it bears emphasis that the asynchronicity of the method is an algorithmic consequence of its formulation, as opposed to the result of further approximations or *a priori* selection of different time steps.

Definition 4.3.1 Suppose $\left(t_k, q_k, \dot{q}_k^+\right)$, the localization matrices $Lo^{(j)}$, and a piecewise-constant approximation of the potential energy $V_h = \sum_j V_h^{(j)}$ are given. Let t_{k+1} and q_{k+1} be the time and points of exit of the rectilinear trajectory $q_k + (t - t_k)\dot{q}_k^+$ from the set $\{V^{(i)}(Lo^{(i)}q) = h^{(i)}\mathbb{Z}\}$. Let $\Delta V^{(i)}$ be the energy jump at q_{k+1} in the direction of advance. The updated velocity is, then,

$$\dot{q}_{k+1}^+ = \dot{q}_k^+ + \lambda_{k+1} M^{-1} n_{k+1} \tag{4.34}$$

where $n_{k+1} = (Lo^{(i)})^T n^{(i)}(Lo^{(i)} q_{k+1})$ and

$$\lambda_{k+1} = \begin{cases} -2\dfrac{\dot{q}_k^+ \cdot n_{k+1}}{n_{k+1}^T M^{-1} n_{k+1}}, \\ \qquad\qquad \text{if } \left(\dot{q}_k^+ \cdot n_{k+1}\right)^2 < 2\Delta V^{(i)} \left(n_{k+1}^T M^{-1} n_{k+1}\right), \\ \dfrac{-\dot{q}_k^+ \cdot n_{k+1} + \operatorname{sign}\left(\Delta V^{(i)}\right) \sqrt{\left(\dot{q}_k^+ \cdot n_{k+1}\right)^2 - 2\Delta V^{(i)} \left(n_{k+1}^T M^{-1} n_{k+1}\right)}}{n_{k+1}^T M^{-1} n_{k+1}}, \\ \qquad\qquad \text{otherwise,} \end{cases} \tag{4.35}$$

It is now clear that we can take advantage of the symmetry considerations that have been discussed in Gonzalez *et al.* [24]. Since the piecewise constant approximation of the additively decomposed potential energy preserves all the symmetries of the local systems exactly, and the discrete trajectories are exact trajectories of a Lagrangian system, we have the following

Theorem 4.3.1 *The asynchronous energy-stepping time-integration scheme is a symplectic-energy-momentum time-reversible integrator with automatic and asynchronous selection of the time step size in each subdomain. In particular, the scheme conserves exactly all the momentum maps of the original local Lagrangians. Moreover, if the process of localization preserves all the symmetries of the global system, asynchronous energy-stepping preserves all the momentum maps of the original global Lagrangian.*

Symmetry or time-reversibility of asynchronous energy-stepping follows directly from the definition of the scheme. The automatic and asynchronous time-step selection properties also follow by construction. In particular, in regions where the localized potential energy gradient $\nabla V^{(j)}$ is steep,

the energy jumps are more closely spaced and the resulting time steps of the subdomain are small. By contrast, if the localized potential energy gradient is small, the resulting time steps of the subdomain are comparatively large. It bears emphasis that what is conserved along the trajectories of the approximate Lagrangian L_h is the *exact*, time-continuous, momentum map of the original Lagrangian L. This is in contrast to discrete variational integrators, which conserve discrete forms of the momentum maps, instead of the exact, time-continuous, momentum maps of the original Lagrangian L. Thus, in particular: if V is invariant under translations then asynchronous energy-stepping conserves the total linear momentum $p_1 + \cdots + p_N$ of the system; and if V is invariant under rotations then asynchronous energy-stepping conserves the total angular momentum $q_1 \times p_1 + \cdots + q_N \times p_N$ of the system. Throughout this chapter we consider configurations $q \in Q = E(n)^N$ and momenta $p \in T_q^*Q$, where $E(n)$ is the Euclidean space of dimension n (i.e., $d = nN$).

4.3.1 Global and local conservation laws

In the context of continuum solid mechanics, the classical theorem of Noether states that each variational symmetry of a Lagrangian $L(\varphi, \dot{\varphi})$ leads to local and global conservation laws, where φ is the deformation mapping [57]. If the system is free of external forces, the *global* momentum map J defined by a symmetry is a constant of the motion, i. e., it remains constant along trajectories. If the conservation law is localized to the subdomain $\Omega^{(s)}$, the *local* momentum map $J^{(s)}$ defined by a symmetry is a constant of motion that additionally accounts for momentum fluxes across subdomain boundaries, i.e., $\partial\Omega^{(s)} \backslash \partial\Omega$.

Likewise, in the context of finite-dimensional Lagrangian systems $L(q, \dot{q})$ (e. g., Lagrangian systems obtained from the finite-element discretization of the action of a continuum solid [51], or Lagrangian systems employed in molecular dynamics simulations), each symmetry of the system leads to local and global conservation laws in accordance with Noether's theorem. We restrict attention to those scenarios where the process of localization preserves all the symmetries of the global system, therefore asynchronous energy-stepping exactly conserves all global momentum maps of the system, as described above, and all local momentum maps as we examine next.

Let G be a Lie group with Lie algebra $\mathfrak{g} = T_eG$. A left action of G on the local configuration space $Q^{(s)}$ is a mapping $\Phi^{(s)} : G \times Q^{(s)} \to Q^{(s)}$, and the infinitesimal generator of $\Phi^{(s)}$ corresponding to $\xi \in \mathfrak{g}$ is the vector field $\xi_{Q^{(s)}} \in TQ^{(s)}$. We say that the local Lagrangian $L^{(s)}$ is invariant under the action $\Phi^{(s)}$ if

$$L^{(s)}(\Phi^{(s)}(g, q^{(s)}), T\Phi^{(s)}(g, q^{(s)})\dot{q}^{(s)}) = L^{(s)}(q^{(s)}, \dot{q}^{(s)}), \quad \forall g \in G,\ (q^{(s)}, \dot{q}^{(s)}) \in TQ^{(s)} \tag{4.36}$$

Then, the local momentum map $J^{(s)} : TQ^{(s)} \to \mathfrak{g}^*$ defined by the action $\Phi^{(s)}$, which expresses a symmetry of $L^{(s)}$, follows from the identity

$$\langle J^{(s)}, \xi \rangle \Big|_0^T = \langle \partial_{\dot{q}^{(s)}} L^{(s)}, \xi_{Q^{(s)}} \rangle \Big|_0^T + \sum_{j \neq s} \left[\langle \partial_{\dot{q}^{(s)}} L^{(j)}, \xi_{Q^{(s)}} \rangle \Big|_0^T - \int_0^T \langle \partial_{q^{(s)}} L^{(j)}, \xi_{Q^{(s)}} \rangle dt \right], \quad \forall \xi \in \mathfrak{g} \tag{4.37}$$

Naturally, the coupling terms in $L^{(j)}$, i.e., terms which involve configurations that belong to $\mathrm{Im}(Lo^{(j)\cap(s)}q^{(s)})$, result in momentum fluxes across subdomain boundaries. Classical examples include:

i) *Local conservation of linear momentum.* In this case, $Q^{(s)} = E(n)^{N_s}$, $G = E(n)$ and $\Phi^{(s)}(u, q^{(s)}) = \{q_1^{(s)} + u, \ldots, q_{N_s}^{(s)} + u\}$ represents a rigid translation of the system by $u \in E(n)$. The corresponding momentum map is the total linear momentum of the local system plus a linear momentum flux,

$$J^{(s)} \Big|_0^T = \sum_{a=1}^{N_s} p_a^{(s)} \Big|_0^T + \sum_{j \neq s} \sum_{a \in \omega^{(j,s)}} \left[p_a^{(j)} \Big|_0^T + \int_0^T \partial_{q_a^{(j)}} V^{(j)} dt \right]$$

where $\omega^{(j,s)} = \{a : q_a^{(j)} \in \mathrm{Im}(Lo^{(j)\cap(s)}q^{(s)})\}$.

ii) *Local conservation of angular momentum.* In this case, $Q^{(s)} = E(n)^{N_s}$, $G = SO(n)$ and $\Phi^{(s)}(R, q^{(s)}) = \{Rq_1^{(s)}, \ldots, Rq_{N_s}^{(s)}\}$ represents a rigid rotation of the system by $R \in SO(n)$. The corresponding momentum map is the total angular momentum of the local system plus an

angular momentum flux,

$$J^{(s)}\Big|_0^T = \sum_{a=1}^{N_s} q_a^{(s)} \times p_a^{(s)}\Big|_0^T + \sum_{j\neq s} \sum_{a\in\omega^{(j,s)}} \left[q_a^{(j)} \times p_a^{(j)}\Big|_0^T + \int_0^T q_a^{(j)} \times \partial_{q_a^{(j)}} V^{(j)} dt \right]$$

iii) *Local conservation of energy.* In this case, a space-time configuration manifold is considered, i.e., $\mathbb{Q}^{(s)} = \mathbb{R} \times Q^{(s)}$, $G = \mathbb{R}$ and $\Phi^{(s)}(u, (t^{(s)}, q^{(s)})) = (t^{(s)} + u, q^{(s)})$ represents a time-shift by $u \in \mathbb{R}$. The corresponding momentum map is the total energy of the local system plus an energy flux,

$$J^{(s)}\Big|_0^T = -\, E^{(s)}\Big|_0^T - \sum_{j\neq s} \sum_{a\in\omega^{(j,s)}} \left[\frac{1}{2} p_a^{(j)} \cdot \dot{q}_a^{(j)}\Big|_0^T + \int_0^T \partial_{q_a^{(j)}} V^{(j)} \cdot \dot{q}_a^{(j)} dt \right]$$

A particularly appealing property of the additive decomposition of the Lagrangian (4.3) assumed in this work and the piecewise approximation of the localized potential energies (4.7) is that they preserve all the symmetries of the original system exactly. To verify this, we simply observe that $V_h^{(j)}$ has all the symmetries of $V^{(j)}$—which itself has all the symmetries of V by assumption—, that is

$$V_h^{(j)} \circ \Phi_g^{(j)} = (h\lfloor h^{-1} V^{(j)} \rfloor) \circ \Phi_g^{(j)} = h\lfloor h^{-1} V^{(j)} \circ \Phi_g^{(j)} \rfloor = h\lfloor h^{-1} V^{(j)} \rfloor = V_h^{(j)}, \quad \forall g \in G \tag{4.38}$$

where G is a symmetry group of $V^{(j)}$ and $\Phi^{(j)}$ is an action that leaves $V^{(j)}$ invariant, i.e., $V^{(j)} \circ \Phi^{(j)} = V^{(j)}$. Then, it follows from the regularization procedure of Theorem 2.2.1 (Theorem 2.1 in [24]), and a standard approximation argument, that the local momentum map $J^{(s)}$ is constant along asynchronous energy-stepping trajectories given by Definition 4.2.1.

4.4 Implementation and computational efficiency

4.4.1 Implementation

The implementation of the asynchronous energy-stepping scheme is summarized in Algorithm 5. The algorithm consists of three methods. The first method SMALLEST-ROOT determines the earliest root of the equation

$$V^{(j)}(q_k^{(j)} + (t_{k+1}^{(j)} - t_k^{(j)})\dot{q}_k^{(j)}) - V^{(j)}(q_k^{(j)}) + \Delta V^{(j)} = 0, \quad \forall j \in [1, n_{LV}] \tag{4.39}$$

where $\Delta V^{(j)}$ can take values in $\{0, h^{(j)}, -h^{(j)}\}$. This task can be effectively accomplished by locating first, by means of an incremental search technique, a time interval containing the smallest positive root and then zeroing in on the root by means of an iterative procedure such as bisection, Newton-Raphson, or a combination thereof. The active subdomain, i.e., the i-th subdomain, and the nature of the step, i.e., whether it consists of a diffraction by an uphill energy step, a diffraction by a downhill energy step or a reflection at an uphill energy step, is identified simultaneously with the computation of $t_{k+1}^{(i)}$. The second method UPDATE-VELOCITIES is responsible for updating velocities according to Definition 4.2.1 and reduces to only two scenarios, as it was the case in the energy-stepping scheme [24]. The third method BOUNDARY-COMMUNICATOR modifies the boundary subdomains with the recently updated positions and velocities.

Algorithm 5 Asynchronous energy-stepping integrator

Require: $\{Lo^{(j)}\}$, $\{V^{(j)}\}$, q_0, $\dot{q}_0$, t_0, t_f and $\{h^{(j)}\}$

$\{k^{(j)}\} \leftarrow \{0\}$
while $\min\{t_k^{(j)}\} < t_f$ **do**
  $t_{k+1}^{(i)} \leftarrow$ SMALLEST-ROOT$\Big(V^{(j)}(q_k^{(j)} + (t_{k+1}^{(j)} - t_k^{(j)})\dot{q}_k^{(j)}) - V^{(j)}(q_k^{(j)}) + \Delta V^{(j)} = 0\Big)$
  $q_{k+1}^{(i)} \leftarrow q_k^{(i)} + (t_{k+1}^{(i)} - t_k^{(i)})\dot{q}_k^{(i)}$
  $n_{k+1}^{(i)} \leftarrow \nabla V^{(i)}(q_{k+1}^{(i)})$
  $\dot{q}_{k+1}^{(i)} \leftarrow$ UPDATE-VELOCITIES$\Big(\dot{q}_k^{(i)}, n_{k+1}^{(i)}, h^{(i)}\Big)$
  $< q_{k+1}^{(j)}, \dot{q}_{k+1}^{(j)} > \leftarrow$ BOUNDARY-COMMUNICATOR$\Big(q_{k+1}^{(i)}, \dot{q}_{k+1}^{(i)}\Big)$
  $k^{(i)} \leftarrow k^{(i)} + 1$
end while

4.4.2 Computational efficiency

In analyzing the computational efficiency of an algorithm relative to others, one usually focuses on speedup. We first recall that the speedup of a parallelized implementation of an algorithm relative to its serial counterpart is defined as

$$\sigma = \frac{T_1}{T_p} \tag{4.40}$$

where T_1 is the sequential execution time on one processor and T_p is the parallel execution time on p processors. Under the assumption that the problem size remains fixed when parallelized, Amdahl's law provides the maximum speedup that can be achieved when a fraction s (with $0 \leq s \leq 1$) of the algorithm remains sequential and only a fraction $1 - s$ is amenable to p-fold parallel execution

$$\sigma = \frac{p}{1 - s + sp} \tag{4.41}$$

Likewise, in order to asses the computational efficiency of asynchronous energy-stepping, we define the speedup of an asynchronous method relative to its synchronous counterpart as

$$\sigma = \frac{T_{\mathrm{ES}}}{T_{\mathrm{AES},p}} \tag{4.42}$$

where T_{ES} is the execution time of (synchronous) energy-stepping and $T_{\mathrm{AES},p}$ is the execution time of asynchronous energy-stepping based on a p-fold decomposition of the domain.

It bears emphasis that the resulting speedup obtained from the application of Definitions (4.40) and (4.42) depends on the problem under consideration. The speedup of a parallelized implementation of an algorithm depends on the fraction s of the algorithm, in terms of the computational time-cost, that remains sequential. Similarly, the speedup of an asynchronous method which is based on a p-fold decomposition of the domain will depend on the dynamics of the problem under consideration. In Section 4.5, we illustrate the speedup of asynchronous energy-stepping under different scenarios.

4.4.3 Domain decomposition

Domain decomposition or mesh partitioning is relevant to the effective and efficient execution of mesh-based scientific simulations on parallel architectures. The strategy one has to follow is such that the computations are balanced and the interprocessor communication is minimized. This can be effectively achieved, among others, by solving a graph-partitioning problem that minimizes the number of edges cut by the decomposition and that keeps an equal number of vertices in each partition. Algorithms for solving such problems are implementing, for example, in the library METIS [44]. We show in Figure 4.2 the finite-element mesh of a cube partitioned by using METIS.

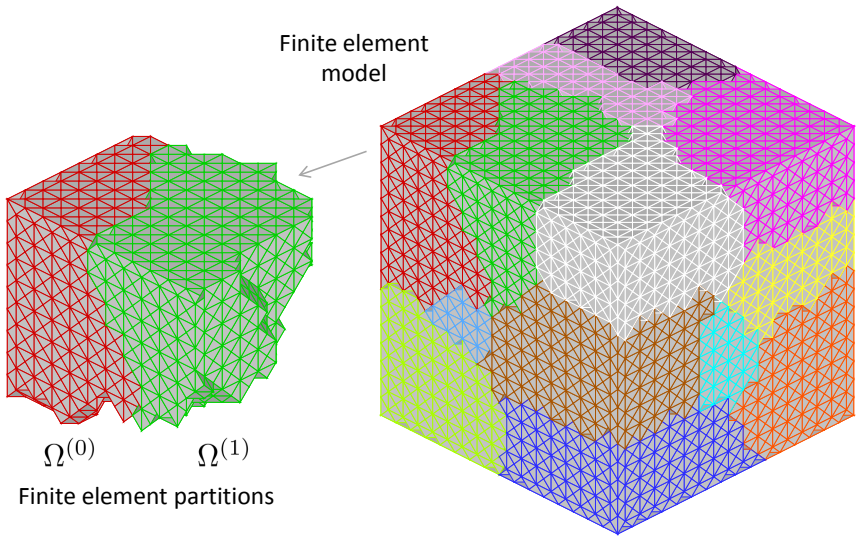


Figure 4.2: Domain decomposition of the finite-element mesh of a cube by recourse to the mesh partitioning library METIS.

The properties of an optimal domain decomposition for asynchronous energy-stepping are not entirely clear in general meshes or domains, if beyond the scope of this work. However, decomposition methods apply in a natural way in particular applications. This is the case of coupled systems with an interface, e. g., many-body contact dynamics, where a naïve strategy one may follow is to identify each subsystem or body in the system as a partition of the domain. We show in Figure 4.3 the mesh partitioning of a finite-element model that comprises the head-on impact of two balls. It is worth noting that three partitions are defined in this example, i.e., each ball results in a finite-element partition and the set of boundary-faces pairs that characterize the interpenetration constraint is

also a primitive subdomain. This natural and intuitive decomposition also appears to be an effective and efficient strategy in the context of asynchronous energy-stepping integrators. In particular, each subdomain will have its own time scale with no detriment to accuracy in the resolution of the fine structure of collisions.

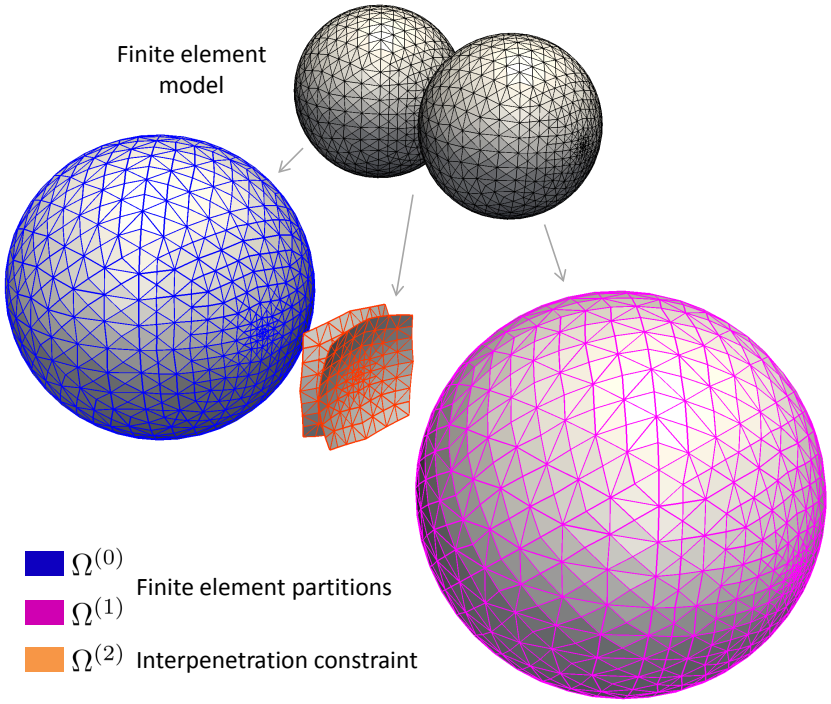


Figure 4.3: Natural and efficient partition of a finite-element model that comprises the head-on impact of two balls.

4.5 Numerical examples

Next, we present selected examples of application which showcase the efficiency, convergence, accuracy, long-term behavior, asynchronicity, automatic time step size selection and conservation properties of asynchronous energy-stepping. We select two areas of application where those properties play an important role. For example, good energy conservation and automatic time size selection are essential in many-body contact dynamics, where the fine structure of collisions may influence significantly the overall behavior of the system. Asynchronous time size selection is furthermore important in reducing the computational cost of—or even in making affordable—many-body dynamics

simulations, where a wide range of strain rates and time scales may coexist in time and space. Good long-term behavior, exact local and global conservation properties, and asynchronous time size selection are likewise relevant to the effective and efficient numerical solution of those problems where the interesting dynamics takes place in a small region of the domain and this region of interest drifts in space as time progresses (e. g., propagation of shock waves, solitary waves, structural and material instabilities, or fracture dynamics). We illustrate the performance of asynchronous energy-stepping in those areas of application by means of two examples: the head-on collision of two soft elastic spherical balls, and the propagation of solitary waves in a chain of elastic beads.

For purposes of assessing the performance of asynchronous energy-stepping, in the subsequent examples we draw detailed comparisons with energy-stepping [24] and with the second-order explicit Newmark method, namely, the member of the Newmark family of time-stepping algorithms corresponding to parameters $\beta = 0$ and $\gamma = 1/2$ (cf., for example [42] for a detailed account of Newmark's method). In the linear regime, explicit Newmark is second-order accurate and conditionally stable, with a critical time step equal to twice over the maximum natural frequency of the system. Detailed analyses of the implicit members of the Newmark family of algorithms, their stability and energy preserving properties (for linear systems) were given in Belytschko and Schoeberle [4], Hughes [35] and related papers.

4.5.1 Head-on collision of soft elastic spherical balls

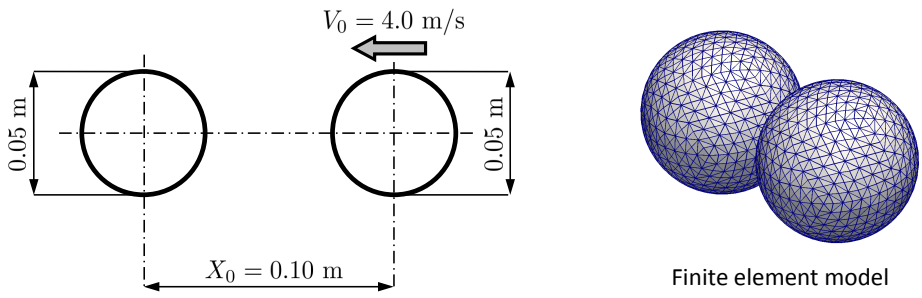


Figure 4.4: Setup of the head-on collision of two soft elastic balls of radius 0.025 m.

We proceed to illustrate the performance of asynchronous energy-stepping in contact dynamics applications by analyzing the head-on collision of two soft elastic spherical balls of radius 0.025 m, Figure 4.4. Initially, the balls are spaced 0.10 m apart, the target ball is stationary and the incoming ball is imparted an initial head-on velocity of magnitude 4.0 m/s. The mesh comprises 34,080 4-node tetrahedral isoparametric elements and 6,994 nodes, and it is partitioned as depicted in Figure 4.3. The balls are compressible neo-Hookean solids characterized by a strain-energy density of the form

$$W(F) = \frac{\lambda_0}{2} (\log J)^2 - \mu_0 \log J + \frac{\mu_0}{2} \operatorname{tr}\left(F^T F\right) \tag{4.43}$$

The values of the material constants in (4.43) are $\lambda_0 = 12.04 \cdot 10^4$ Pa, $\mu_0 = 1.34 \cdot 10^4$ Pa, and the density is $\rho = 1059$ kg/m^3. The material properties are those of an acrylamide gel synthesized by a standard polymerization in water (100 g) of acryamide monomer (10 g) and cross-linker of acrylamide (0.30 g)—rheological characterization of this gel revealed that it behaves as an elastic material in the frequency range from 0.1 to 100 Hz [77]. A sequence of snapshots of the asynchronous energy-stepping trajectory of the target ball corresponding to $h = 5 \cdot 10^{-5}$ J are shown in the bottom row of Figure 4.5. The large deformation undergone by the gel ball is evident in the figure.

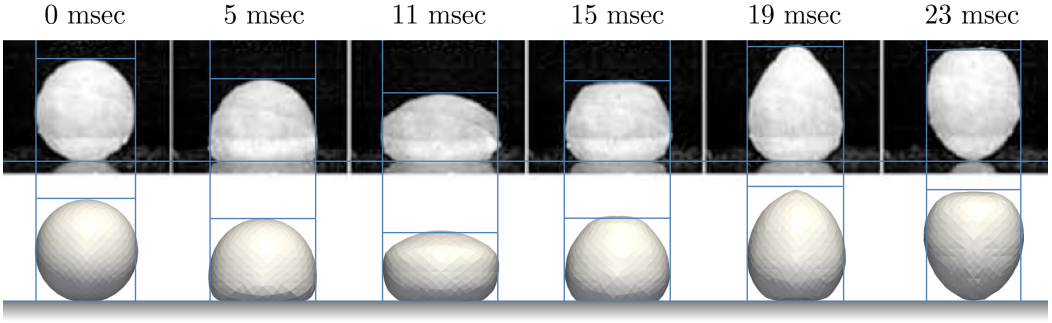


Figure 4.5: Head-on collision of two soft elastic balls of radius 0.025 m. The bottom row shows snapshots of the asynchronous energy-stepping trajectory with $h = 5 \cdot 10^{-5}$ J. The top row shows high-speed camera snapshots of the deformation process of an acrylamide gel ball of radius 0.025 m that impacts a rigid wall at a head-on velocity of magnitude 2.0 m/s (figure adapted from [77]). The first, third, and last pictures show the gel at the initial contact, the maximal deformation and the taking-off, respectively.

Tanaka and co-workers [75–77] have performed experimental studies of the impact of spherical acrylamide gel balls of radius 0.025 m on a rigid substrate at high impact velocities, i.e., beyond the

Hertzian regime. A sequence of high-speed video camera snapshots of the deformation process of a gel ball, with composition and size described above, impacting a rigid wall at a head-on velocity of magnitude 2.0 m/s [77] are shown in the top row of Figure 4.5. Clearly, the numerical setup described above is equivalent to the experimental setup of Tanaka and co-workers. Furthermore, it is interesting to observe that the agreement between asynchronous energy-stepping trajectories and experimental observations is excellent.

The time histories of the total energy, potential energy and kinetic energy of the asynchronous energy-stepping solution corresponding to $h = 5 \cdot 10^{-5}$ J are shown in Figure 4.6. As previously illustrated in Gonzalez *at al.* [24], energy-stepping is observed to conserve energy exactly throughout collisions. Additionally, asynchronous energy-stepping exhibits exact local conservation properties discussed in Section 4.3. Then, before and after the collision, the total energy of each subdomain, i.e., each gel ball, is exactly preserved by asynchronous energy-stepping, as illustrated in Figure 4.7. Through the collision, linear momentum and energy are transfered from incoming to target balls and the method automatically modulates the time-step to resolve the fine structure of the intricate interactions that occur during this process. Figure 4.8 shows time histories of the linear momentum of the system and of each individual subdomain, i.e., incoming and target balls. We again emphasize that asynchronous energy-stepping accounts for the exact momentum fluxes across subdomian boundaries, through the collision, and it exactly preserves all the momentum maps of each subdomain, otherwhile (Section 4.3).

We now illustrate the ability of asynchronous energy-stepping in particular, and energy-stepping integrators in general, to capture the intricate interactions that occur through the collision. The contact force is a footprint of these dynamic interactions that is amenable to experimental validation. Tanaka and co-workers [77] have studied the profile of the contact force during the impact of spherical acrylamide gel balls. Then, in order to validate the model, the contact force is evaluated by simply applying Newton's second law, that is the external force acting on the s-th subdomain is given by

$$F_{ext}^{(s)} = \frac{d}{dt} J^{(s)} \tag{4.44}$$

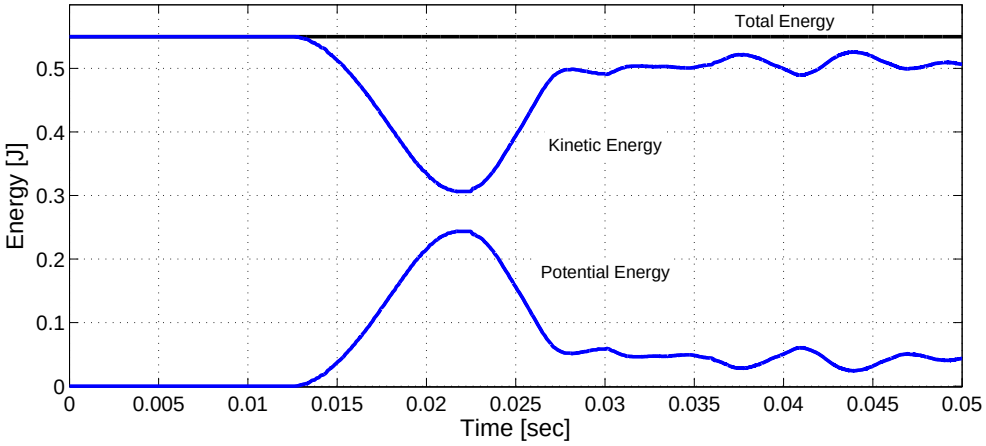


Figure 4.6: Head-on collision of two soft elastic balls of radius 0.025 m. Asynchronous energy-stepping solution with $h = 5 \cdot 10^{-5}$ J. Time history of total, kinetic and potential energies.

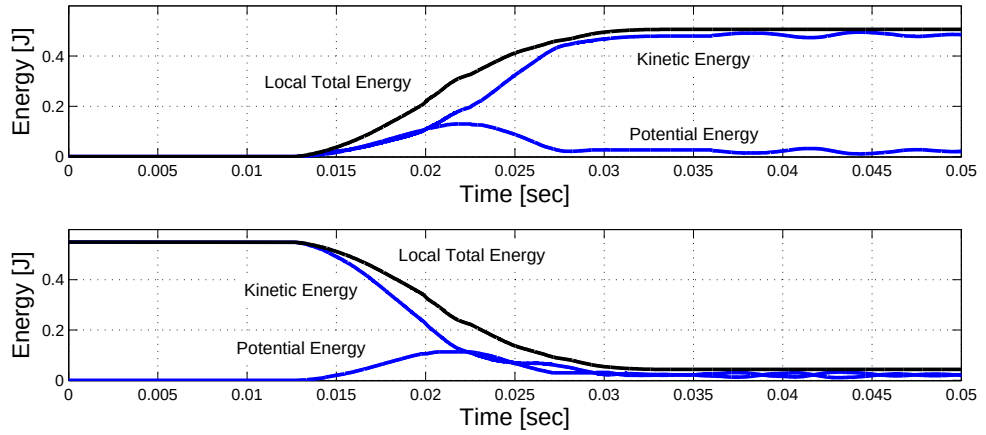


Figure 4.7: Head-on collision of two soft elastic balls of radius 0.025 m. Asynchronous energy-stepping solution with $h = 5 \cdot 10^{-5}$ J. The top and bottom figures correspond to the target and incoming balls, respectively.

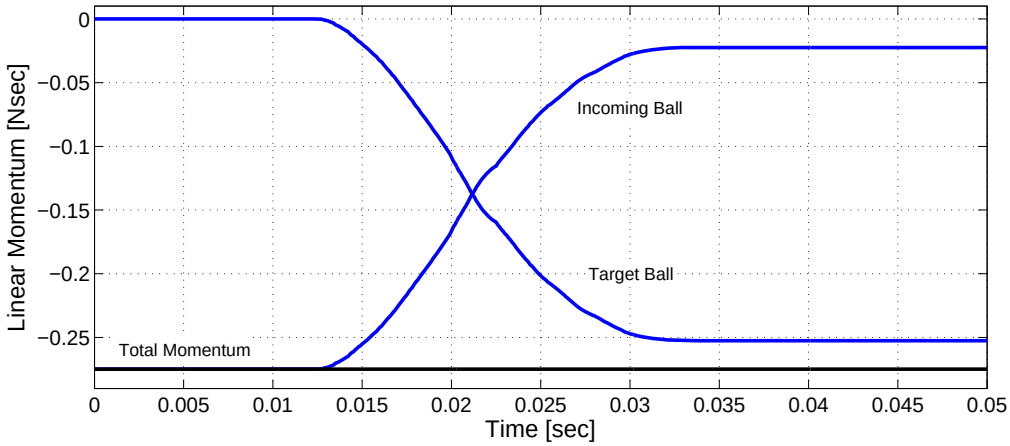


Figure 4.8: Head-on collision of two soft elastic balls of radius 0.025 m. Asynchronous energy-stepping solution with $h = 5 \cdot 10^{-5}$ J. Conservation of total linear momentum and time history of individual subdomains—only the component align with the direction of impact is shown.

where $J^{(s)}$ is the linear momentum of the subdomain, i.e., $J^{(s)} = \sum_{a=1}^{N_s} p_a^{(s)}$. Figure 4.9 shows experimental observations and numerical predictions of the contact force that corresponds to the setup discussed above. The good agreement of the model is evident in the figure.

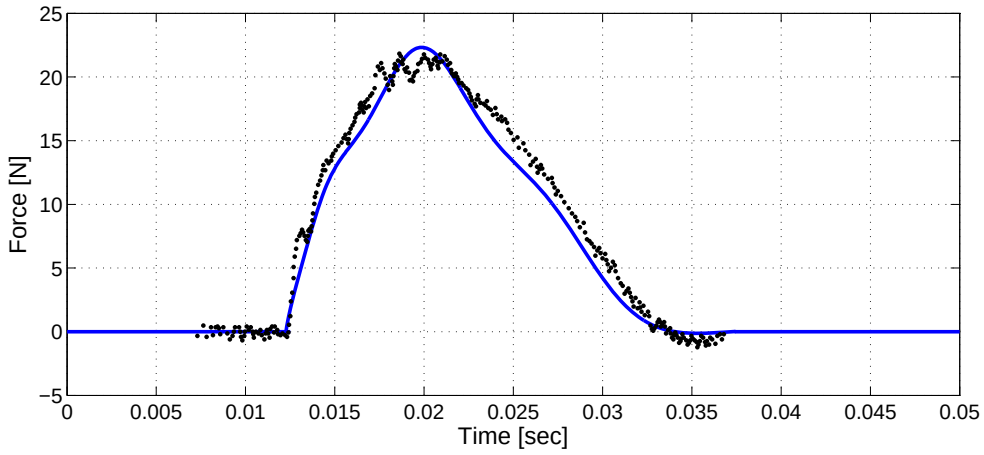


Figure 4.9: Head-on collision of two soft elastic balls of radius 0.025 m. The curve corresponds to the contact force computed from the asynchronous energy-stepping solution with $h = 5 \cdot 10^{-5}$ J. The dots correspond to the contact force measured during the deformation process of an acrylamide gel ball of radius 0.025 m that impacts a rigid wall at a head-on velocity of magnitude 2.0 m/s [77].

A simple modification of the previous setup facilitates the analysis of a different aspect of asynchronous energy-stepping performance. We now consider the head-on impact of two soft elastic balls

of radius 0.025 m and 0.050 m, Figure 4.10. The difference in size of the balls results in one body suffering larger deformations than the other. Then, each ball deforms with different strain rates and time scales, though they are strongly coupled by the fine structure of the collision which has to be resolved accurately. The model has the same mesh topology, strain-energy density and material constants as the previous setup. A sequence of snapshots of the asynchronous energy-stepping trajectory corresponding to $h = 5 \cdot 10^{-5}$ J are shown in Figure 4.11. Cross-sections of the mesh at time of contact and at later times during the collision are shown in Figure 4.12. The asymmetry of the large deformations undergone by the gel balls is evident in the figures.

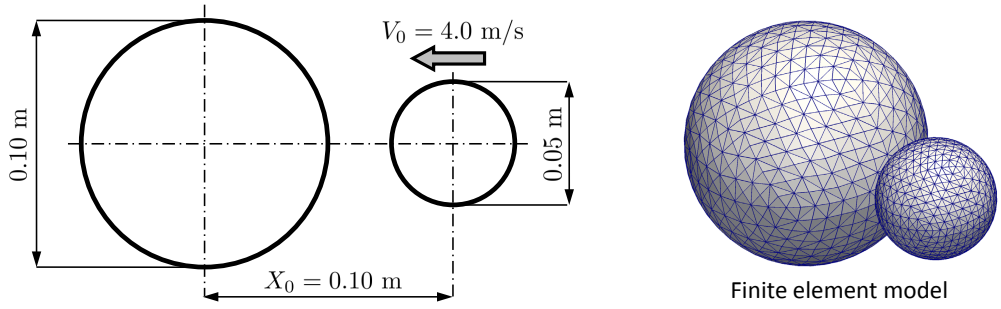
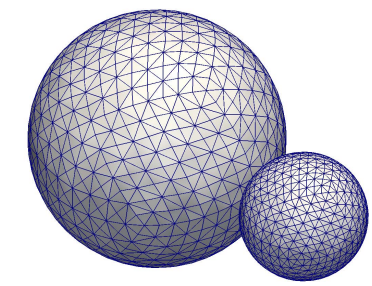
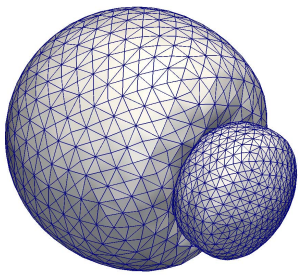


Figure 4.10: Setup of the head-on collision of two soft elastic balls of radius 0.025 m and 0.050 m.

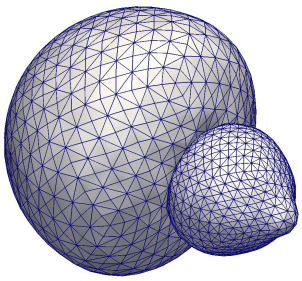
The time histories of global total energy, potential energy and kinetic energy of the asynchronous energy-stepping solution corresponding to $h = 5 \cdot 10^{-5}$ J are shown in Figure 4.13, and their local counterparts are shown in Figure 4.14. It is worth noting the natural asynchrony and different time scale of incoming and target ball motions. We also show in Figure 4.15 the time histories of the global linear momentum and of each local individual contribution.



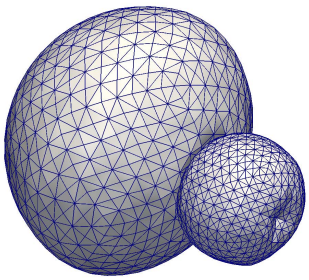
Time: 0 msec + t_C



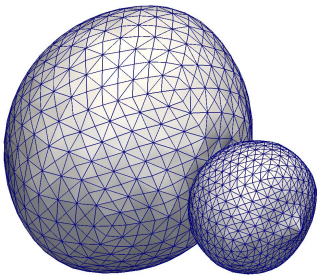
Time: 9 msec + t_C



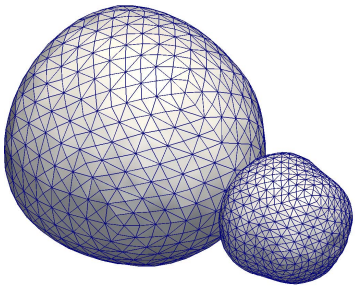
Time: 15 msec + t_C



Time: 17.5 msec + t_C

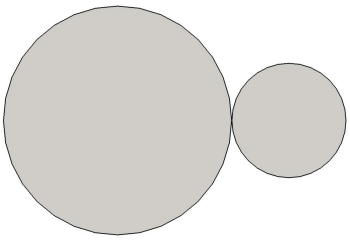


Time: 20 msec + t_C

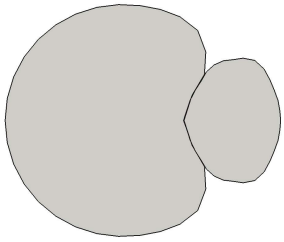


Time: 26 msec + t_C

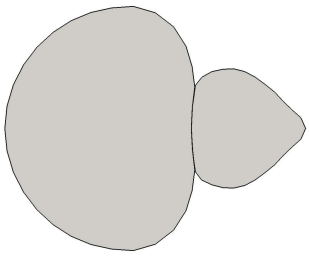
Figure 4.11: Head-on collision of two soft elastic balls of radius 0.025 m and 0.050 m. Snapshots of the asynchronous energy-stepping trajectory with $h = 5 \cdot 10^{-5}$ J, at time of contact t_C and later times during the collision.



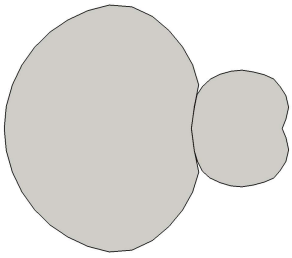
Time: 0 msec + t_C



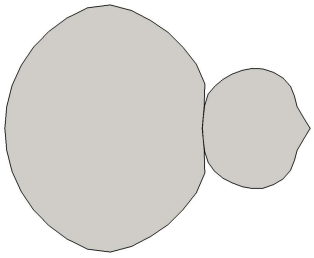
Time: 9 msec + t_C



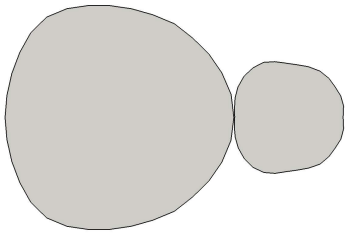
Time: 15 msec + t_C



Time: 17.5 msec + t_C



Time: 20 msec + t_C



Time: 26 msec + t_C

Figure 4.12: Head-on collision of two soft elastic balls of radius 0.025 m and 0.050 m. Cross-sections of the mesh for the asynchronous energy-stepping trajectory with $h = 5 \cdot 10^{-5}$ J, at time of contact t_C and later times during the collision.

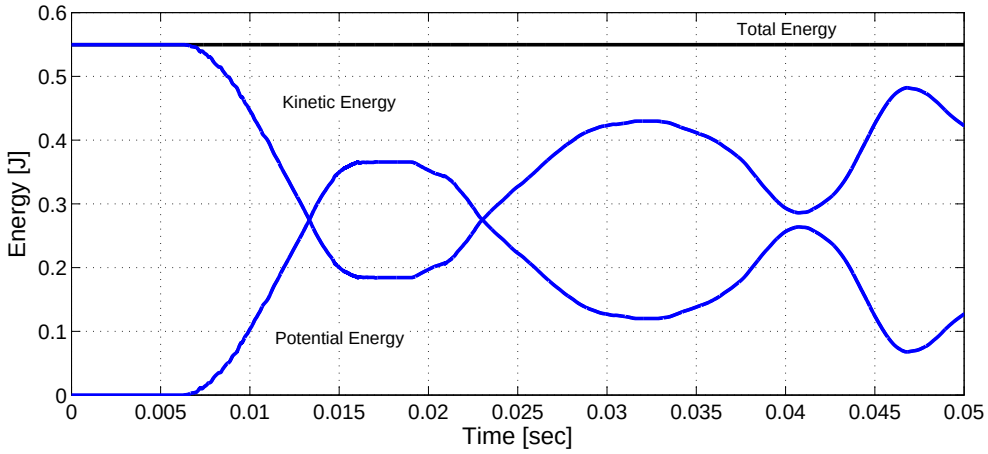


Figure 4.13: Head-on collision of two soft elastic balls of radius 0.025 m and 0.050 m. Asynchronous energy-stepping solution with $h = 5 \cdot 10^{-5}$ J. Time history of total, kinetic and potential energies.

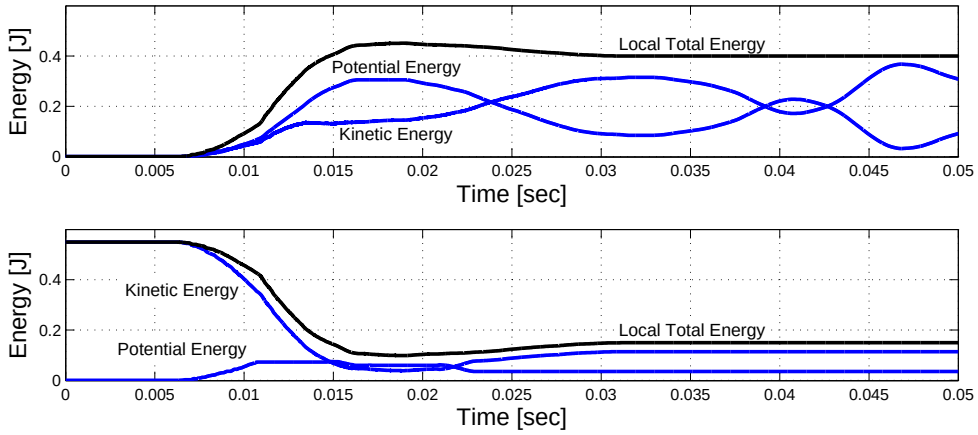


Figure 4.14: Head-on collision of two soft elastic balls of radius 0.025 m and 0.050 m. Asynchronous energy-stepping solution with $h = 5 \cdot 10^{-5}$ J. The top and bottom figures correspond to the target and incoming balls, respectively.

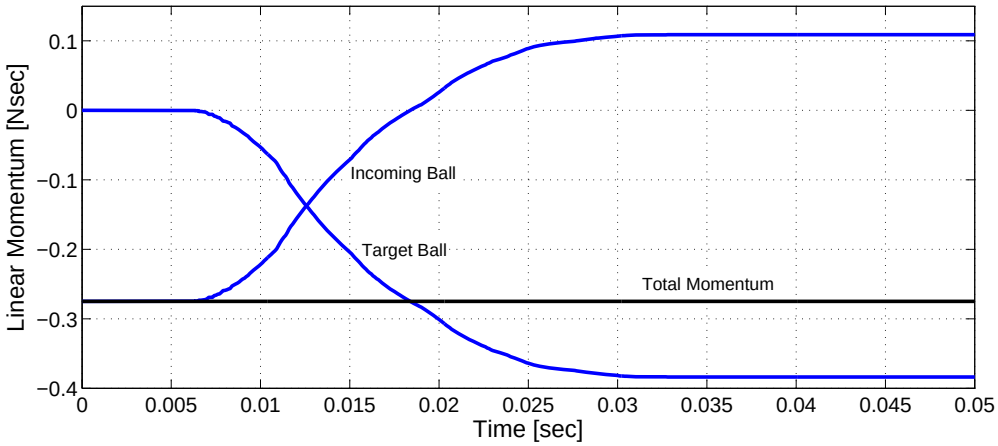


Figure 4.15: Head-on collision of two soft elastic balls of radius 0.025 m and 0.050 m. Asynchronous energy-stepping solution with $h = 5 \cdot 10^{-5}$ J. Conservation of total linear momentum and time history of individual subdomains—only the component align with the direction of impact is shown.

The histograms of the time steps selected by asynchronous energy-stepping in the first (Figure 4.4) and second (Figure 4.10) setup are shown in Figures 4.16 and 4.17, respectively. The broad range of time steps asynchronous energy-stepping employed in each partition and contact constraint is noteworthy. It also bears emphasis the differences and similarities of both setups. While the histograms that correspond to the contact constraint are almost equal for both setups, the histograms that correspond to the incoming and target balls are remarkably different. In the fist setup, the symmetry in the time steps selected by asynchronous energy-stepping suggests that the method does not break the characteristic synchronic behavior of the problem, albeit possibly at the expense of a small computational overhead. By contrast, in the second setup, the asymmetry in the time steps selected by the time-integration scheme clearly shows the automatic and asynchronous time-step selection properties of the method. These properties may result in a speedup of asynchronous energy-stepping over (synchronous) energy-stepping, as we will discuss in our next example.

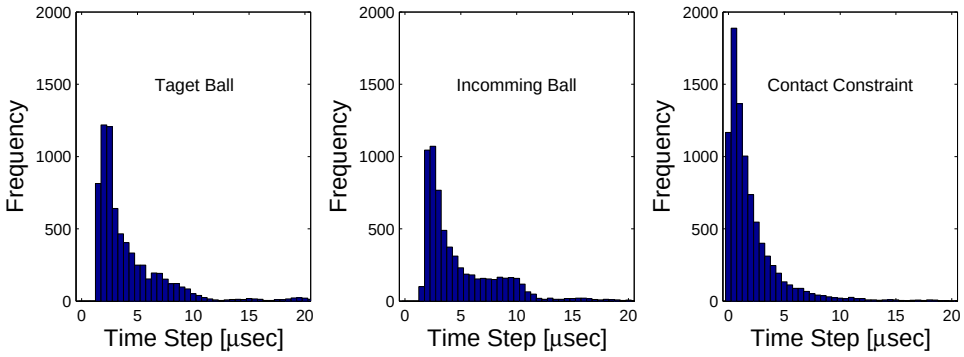


Figure 4.16: Head-on collision of two soft elastic balls of radius 0.025 m. Histogram of time steps selected by asynchronous energy-stepping for $h = 5 \cdot 10^{-5}$ J.

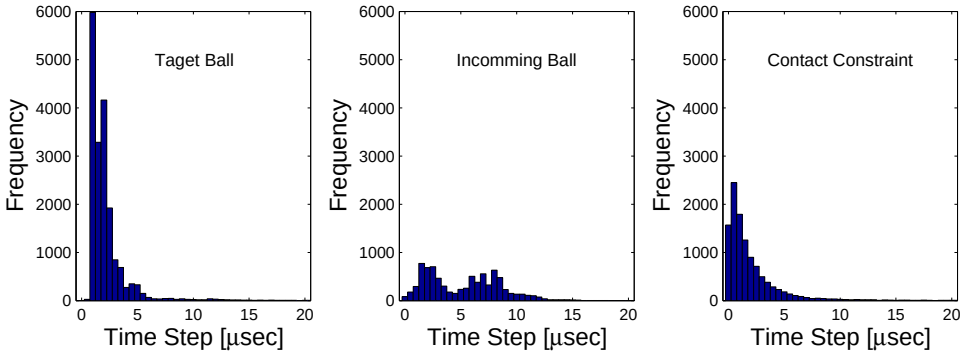


Figure 4.17: Head-on collision of two soft elastic balls of radius 0.025 m and 0.050 m. Histogram of time steps selected by asynchronous energy-stepping for $h = 5 \cdot 10^{-5}$ J.

4.5.2 Chain of elastic beads

Our next example concerns the formation and propagation of solitary waves in a chain of ten elastic beads of diameter 3/8". Initially, a chain of nine beads is stationary and the incident solitary pulse is generated by a striker bead with impact velocity of 0.44 m/s, Figure 4.18. In this example, the interesting dynamics takes place in a small region of the system, i.e., the beads that support the solitary pulse, and this region of interest drifts in space as the wave travels along the chain. The mesh comprises 28,480 4-node tetrahedral isoparametric elements and 6,510 nodes. The balls are compressible neo-Hookean solids characterized by a strain-energy density of the form (4.43).

The values of the material constants correspond to steel and are equal to $\lambda_0 = 9.70 \cdot 10^{10}$ Pa, $\mu_0 = 8.26 \cdot 10^{10}$ Pa, the density is $\rho = 7780$ kg/m^3.

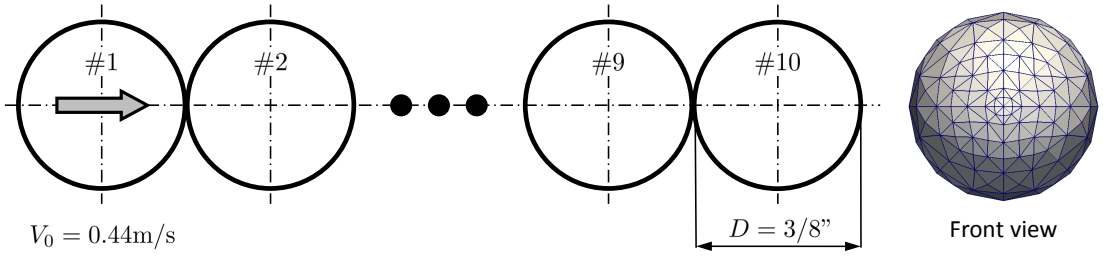


Figure 4.18: Setup of the chain of ten elastic beads with diameter 3/8” and impact velocity of 0.44 m/s. On the right hand side, it is shown a detail of the finite-element mesh employed in the calculations.

The time histories of the total energy, potential energy and kinetic energy of the asynchronous energy-stepping solution corresponding to $h = 10^{-7}$ J are shown in Figure 4.19. The localized nature of the dynamics is clearly observed in the time histories of the local—as opposed to global—energies depicted in Figure 4.20 for the first four beads of the chain. It is also worth noting that, before and after the solitary wave travels along a bead, the local energy is exactly preserved by asynchronous energy-stepping. Similarly, all global momentum maps of the system and all local momentum fluxes across bead boundaries are exactly preserved by asynchronous energy-stepping. In particular, Figure 4.21 shows the exact conservation of total linear momentum and the change in shape of the local contributions $J^{(s)}$ as the solitary wave propagates along the chain.

As we have discussed previously, there are two unique properties of asynchronous energy-stepping: i) automatic and asynchronous selection of the time step size in each subdomain, and ii) exact conservation of all local and global momentum maps of the systems. These features allow for a confident and straightforward analysis of the local behavior of a Lagrangian system. For example, the formation and propagation of solitary waves are of particular interest to the dynamics of a one-dimensional chain of elastic beads. Numerous experimental and theoretical investigations have been conducted with the aim of understanding the shape and the speed of these solitary waves (see, for

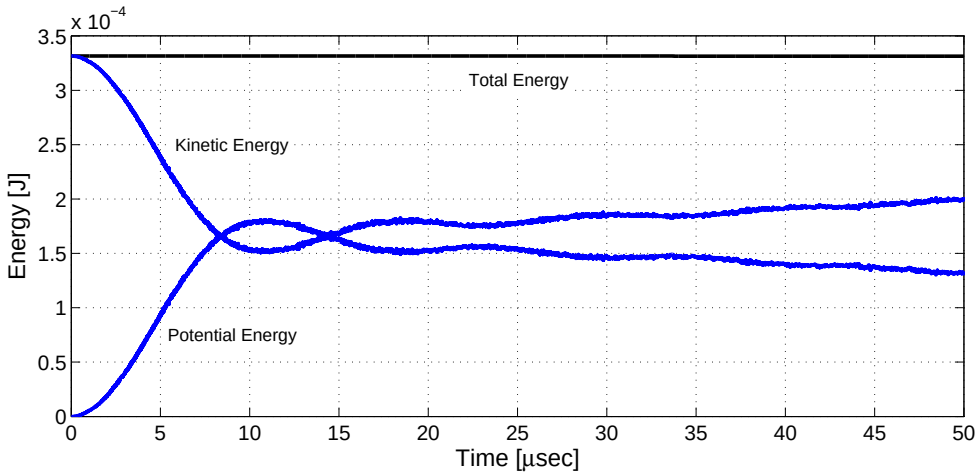


Figure 4.19: Chain of ten elastic beads. Asynchronous energy-stepping solution with $h = 10^{-7}$ J. Time history of total, kinetic and potential energies.

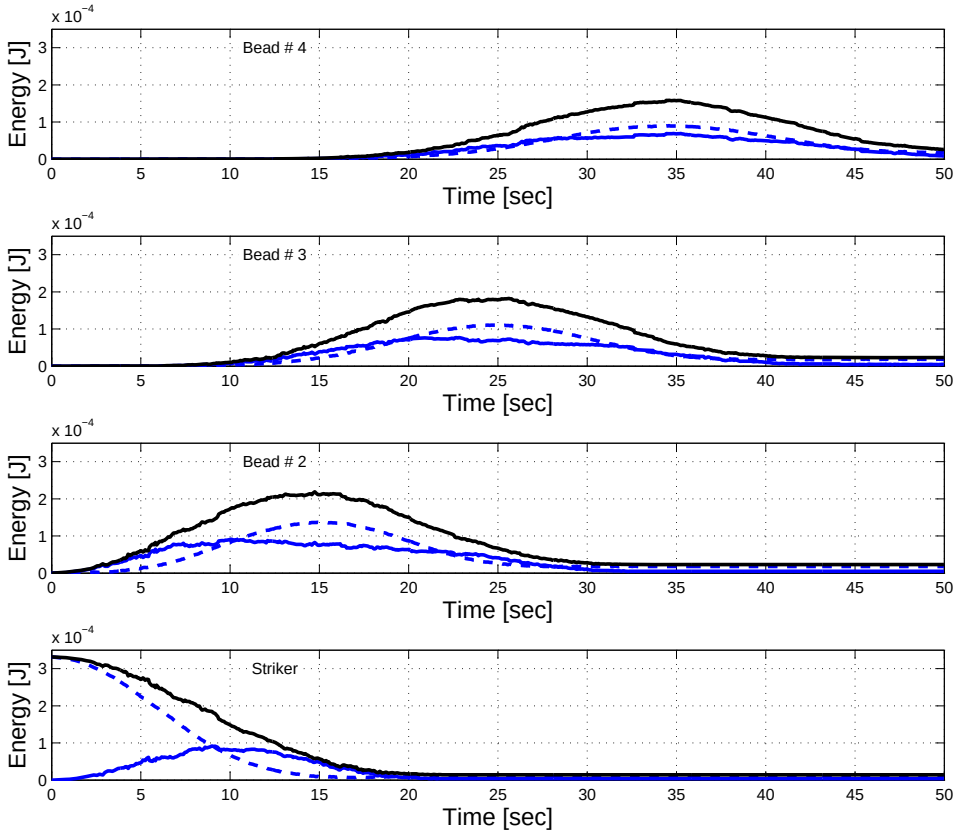


Figure 4.20: Chain of ten elastic beads. Asynchronous energy-stepping solution with $h = 10^{-7}$ J. Time history of total (black curve), kinetic (dashed blue curve) and potential (blue curve) *local* energies.

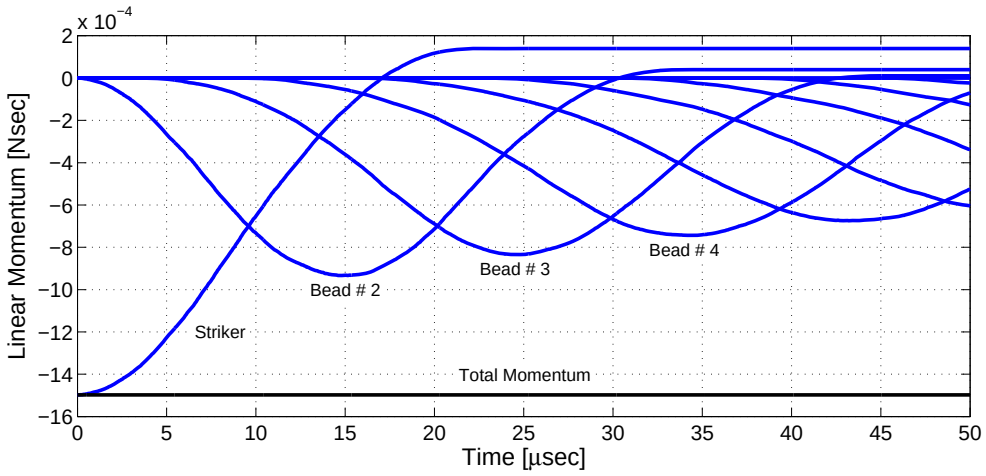


Figure 4.21: Chain of ten elastic beads. Asynchronous energy-stepping solution with $h = 10^{-7}$ J. Conservation of total linear momentum and time history of individual $J^{(s)}$—only the component align with the chain of beads is shown.

example, [9] and references therein). However, the amount of vibrational kinetic energy retained in each bead during and after collision has never been accounted for. This is clearly a local behavior of the system that can be readily studied with asynchronous energy-stepping. The vibrational kinetic energy of each bead is defined as

$$\mathrm{VKE}^{(s)} = \sum_{a=1}^{N_s} \frac{1}{2} m_a \|\dot{q}_a - \langle \dot{q} \rangle^{(s)}\|^2 \tag{4.45}$$

where $\langle \dot{q} \rangle^{(s)}$ is the center of mass velocity of the s-th bead. Figure 4.22 illustrates the amount of vibrational kinetic energy in the system as the solitary wave travels along the chain and it also shows the individual contribution of each bead. It is worth noting that, for this particular finite-element mesh and initial conditions, the vibrational kinetic energy *trapped* in the system is not negligible and naturally accumulates as the solitary wave travels along the chain.

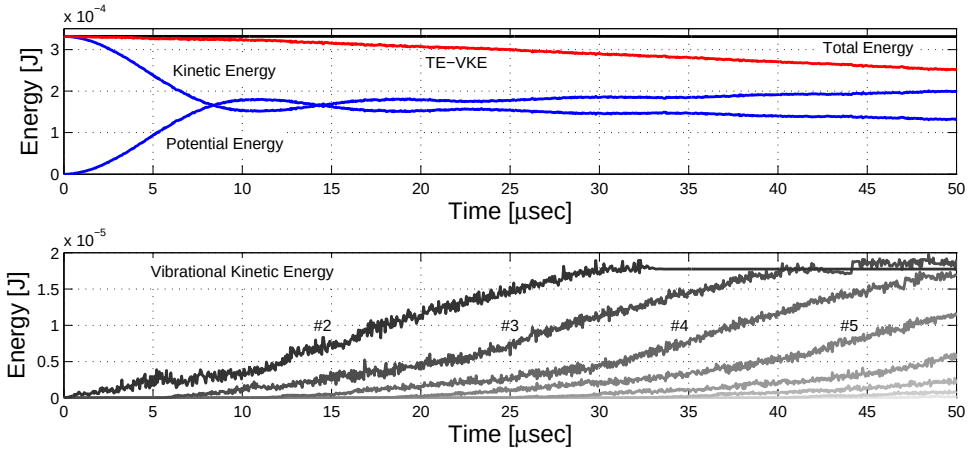


Figure 4.22: Chain of ten elastic beads. Asynchronous energy-stepping solution with $h = 10^{-7}$ J. The top figure illustrates the amount of vibrational kinetic energy (VKE) in the system as the solitary wave travels along the chain. The bottom figure shows the time history of individual components $\mathrm{VKE}^{(s)}$.

Finally, we investigate the computational efficiency of asynchronous energy-stepping over energy-stepping and explicit Newmark for the setup depicted in Figure 4.18. We compare the execution time of asynchronous energy-stepping for $h = 10^{-6}$ J and the mesh partitioned into 10 subdomains ($T_{\text{AES-10},10^{-6}}$), with execution times of energy-stepping for $h = 10^{-6}$ J ($T_{\text{ES},10^{-6}}$) and explicit Newmark for a small time step $\Delta t = 10^{-4}$ μsec ($T_{\text{NM},10^{-4}}$)—this time step is 10 times smaller than the average time step resulting from the energy-stepping calculations. As each numerical time integrator provides a different approximation of the exact solution of the problem, it is not only important to compute the timing of each simulation but also to quantify the accuracy of the numerical approximation. To this end we additionally compute the asynchronous energy-stepping solution for $h = 10^{-7}$ J and we compare the time history of energies for each numerical solution as an indication of accuracy. Figure 4.23 shows that the explicit Newmark solution quickly loses stability and blows up, even though the time step employed is small and therefore the execution time is 25 times larger than $T_{\text{AES-10},10^{-6}}$. Figure 4.24 shows that asynchronous energy-stepping solutions with $h = 10^{-6}$ J and with $h = 10^{-7}$ J are in reasonable agreement. In contrast, Figure 4.25 shows that the energy-

stepping solution is less accurate than its asynchronous counterpart for the same energy step. The execution time for these different cases is summarized in Table 4.1 in terms of speedup relative to $T_{\text{AES-10,10}^{-6}}$ (see Section 4.4.2). It bears emphasis that the asynchronous energy-stepping solution is more computationally efficient and more accurate than its synchronous counterpart. These results suggest that, for this particular finite-element mesh and initial conditions, a speedup of the form

$$\sigma = \frac{T_{\text{ES,10}^{-7}}}{T_{\text{AES-10,10}^{-6}}} \sim 5.2$$

is more appropriate than $\sigma = T_{\text{ES,10}^{-6}}/T_{\text{AES-10,10}^{-6}} \sim 2.5$, as shown in Figure 4.26. Nonetheless, a speedup of ~ 2.5 is consistent with the fact that, for this example, the solitary wave has a support of ~ 4 beads and the chain comprises 10 beads.

	$T_{\text{AES-10,10}^{-6}}$	$T_{\text{ES,10}^{-6}}$	$T_{\text{NM,10}^{-4}}$	$T_{\text{AES-10,10}^{-7}}$	$T_{\text{ES,10}^{-7}}$
Speedup	1	~ 2.5	~ 25.0	~ 1.8	~ 5.2

Table 4.1: Chain of ten elastic beads. Speedup is defined as $T/T_{\text{AES-10,10}^{-6}}$.

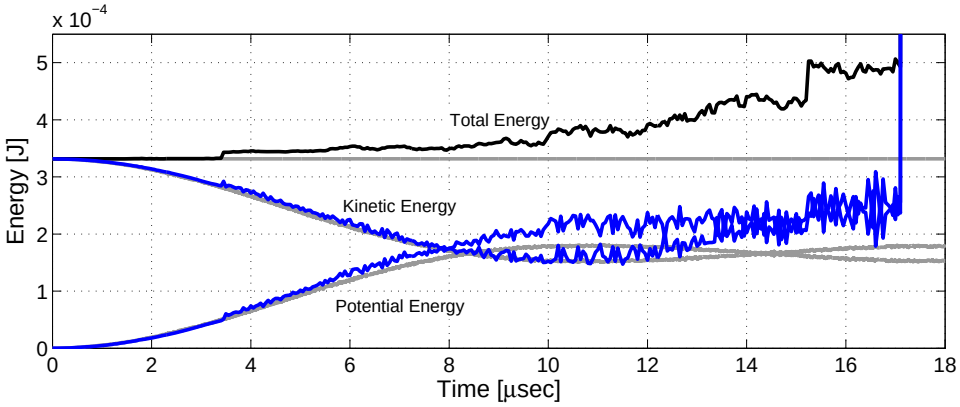


Figure 4.23: Chain of ten elastic beads. Explicit Newmark solution with $\Delta t = 10^{-4}$ μsec. Time history of total, kinetic and potential energies— grey curves correspond to the asynchronous energy-stepping solution with $h = 10^{-7}$ J.

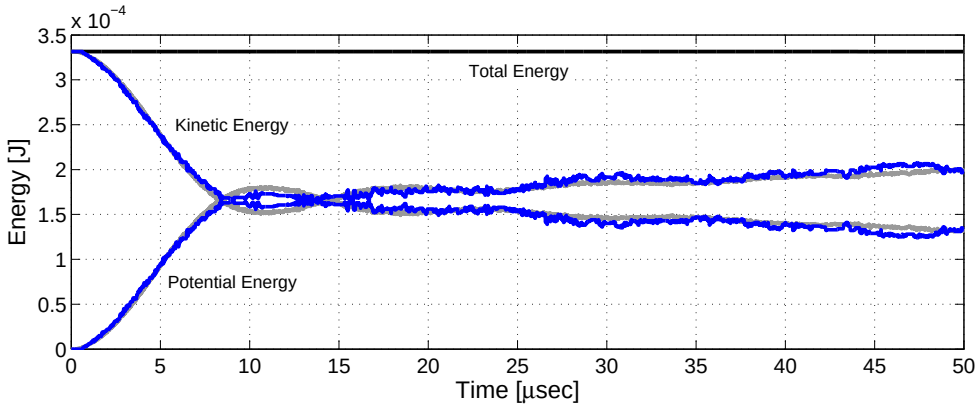


Figure 4.24: Chain of ten elastic beads. Asynchronous energy-stepping solution with $h = 10^{-6}$ J. Time history of total, kinetic and potential energies— grey curves correspond to the asynchronous energy-stepping solution with $h = 10^{-7}$ J.

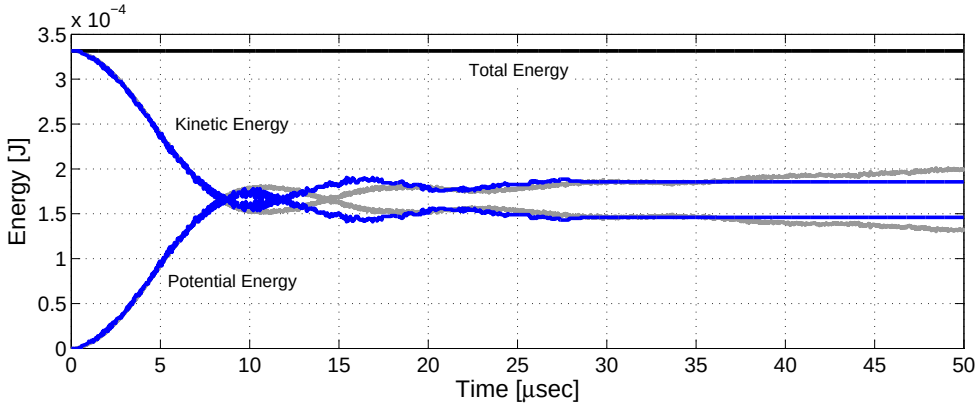


Figure 4.25: Chain of ten elastic beads. Energy-stepping solution with $h = 10^{-6}$ J. Time history of total, kinetic and potential energies— grey curves correspond to the asynchronous energy-stepping solution with $h = 10^{-7}$ J.

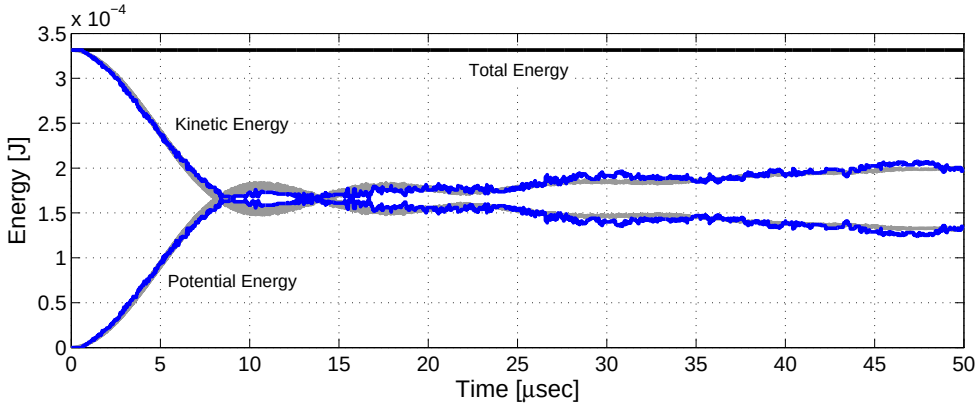


Figure 4.26: Chain of ten elastic beads. Asynchronous energy-stepping solution with $h = 10^{-6}$ J. Time history of total, kinetic and potential energies— grey curves correspond to the energy-stepping solution with $h = 10^{-7}$ J.

4.6 Summary and discussion

We have formulated a new class of time-integration schemes for Lagrangian mechanics, which we refer to as *asynchronous energy-stepping*, that are energy-momentum conserving, symplectic, and convergent with automatic and asynchronous selection of the time step. In order to achieve these properties we introduce a localization of the configuration space that allows for an additive decomposition of the *global* Lagrangian into *local* Lagrangians amenable to energy-stepping approximations. Thus, we replace the original *localized* potential energies by stepwise or *terraced* approximations at steps of uniform heights—energy-stepping integrators in contrast result from stepwise approximations of the *global* potential energy of the system. The trajectories of the resulting Lagrangians can be characterized explicitly and consist of intervals of rectilinear motion that span consecutive level contours of the localized potential energies. The durations of these intervals differ in each subdomain and are regarded as local time steps, whose asynchronic determination is part of the solution process. We have shown that the asynchronous energy-stepping trajectories are symplectic, exactly conserve all the momentum maps of the original system—if process of localization preserves all the symmetries of the global system—and account for the exact momentum fluxes across subdomain boundaries. We have specifically investigated many-body contact problems where localization is the result of identifying each body in the system as a partition of the domain. Selected examples of application, including the head-on collision of soft elastic spherical balls and the propagation of solitary waves in a chain of elastic beads, demonstrate the excellent long-term behavior of asynchronous energy-stepping, its asynchronous and automatic time-step selection property, and the ease with which it deals with many-body concurrent contact dynamics.

We close by pointing out some limitations of our analysis and possible avenues for extensions of the approach.

Firstly, our experience with selected numerical tests appears to indicate that for a given energy step, asynchronous energy-stepping trajectories are more accurate than (synchronous) energy-stepping trajectories. A rigorous local convergence analysis, such as the study of local bounds on the discretization error provided by Gronwall's inequality, could give some insight into this behavior, if

beyond the scope of this work.

Secondly, we have only investigated many-body contact problems where localization results from naïvely identifying each body in the system as a partition of the domain. However, the properties of an optimal domain decomposition for asynchronous energy-stepping are not entirely clear in general. The investigation of optimal localization strategies in the context of other problems (e. g., propagation of shock waves, structural and material instabilities, or fracture dynamics) is a worthwhile direction of future research.

Chapter 5

Mesoscopic approach to granular crystal dynamics

We present a mesoscopic approach to granular crystal dynamics, which comprises a three-dimensional finite-element model and a one-dimensional regularized contact model. The approach aims at investigating the role of vibrational-energy trapping effects in the overall dynamic behavior of one-dimensional chains under small to moderate impact velocities, and it has as only inputs the geometry and the elastic material properties of the individual particles that form the system. The three-dimensional finite-element model resolves the fine mesoscale structure of dynamic collisions and explicitly accounts for the vibrational kinetic energy retained in each bead as a solitary wave propagates along the chain. The one-dimensional regularized contact model accounts for mesoscopic dynamic effects by means of a restitution coefficient, i.e., the vibrational energy that remains trapped after impact is subsumed under the concept of a coefficient of restitution. We present detailed verification and validation results, and extract conclusions about the qualitative and quantitative predictions of this mesoscopic approach.

5.1 Introduction

Granular crystals or highly packed granular lattices, for example one-dimensional chains of beads, are strongly nonlinear systems that exhibit unique wave dynamics. In particular, a chain of elastic spherical beads supports the formation of solitary waves as the result of Hertzian nonlinear contact interactions between the particles in the system [65]. This is in sharp contrast to disordered granular media where the nature of the system is additionally driven by frictional and rotational dynamics. Granular crystals can achieve extremely tunable properties by the simple expedient of combining different materials and sizes, and by the application of weak precompression to the system [14,16,65]. Over the last decade, the response of these systems has drawn considerable attention and many potential applications have been studied, such as sound [conference], shock [17,31] and energy [17,33] absorbing layers, actuators [46], and sound focusing devices [74].

The formation of solitary waves in a chain of elastic spherical beads was discovered analytically and numerically by Nesterenko in 1983 [64] and it was experimentally observed for the first time later in 1985 [49]. Since then, numerous experimental and theoretical investigations have been conducted, leading to a more profound understanding of these strongly nonlinear systems. A typical experimental setup commonly used in these studies consists in a monodisperse one-dimensional chain of spherical beads mounted in a cylindrical guide or arranged in a square four-rod stand. Selected beads are embedded with calibrated piezosensors [15] that allow for detailed measurements of forces acting inside the particles together with the solitary wave speed and duration. Figure 5.1 illustrates a set of experimental observations for a chain composed of 70 steel beads [9]. The characteristic decay of the measured force as the solitary wave propagates along the chain is shown on the left side of the figure. The relationship between solitary wave speed V_s and maximal measured force F_m is plotted on the right side of Figure 5.1. The solid line represents the best linear fit, which gives $V_s \propto F_m^{0.17}$.

A particle mechanics approach has been extensively used in the literature to address modeling and simulation of granular crystal dynamics. In this approach, a number of hypotheses and assumptions are introduced [65]. Firstly, dissipative effects are neglected. Secondly, the interaction

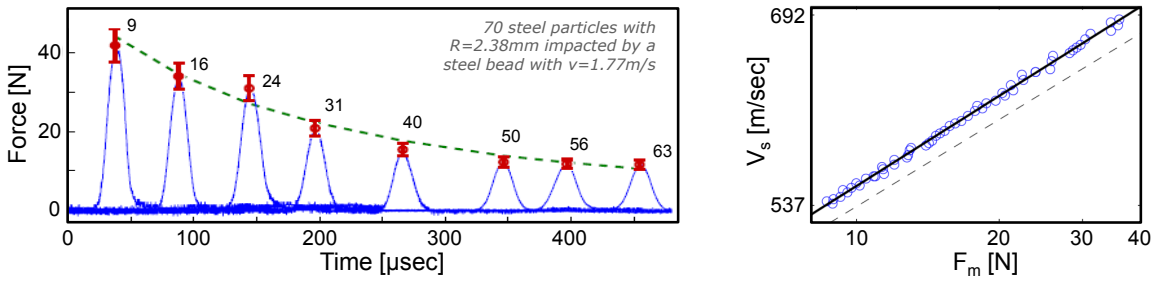


Figure 5.1: Set of experimental observations for a one-dimensional chain composed of 70 steel beads with diameter $D = 0.00476$ m and impacted by a steel bead with $v_{\text{imp}} = 1.77$ m/s. The left side shows the time history of the forces measured by eight sensors. The right side shows the relationship between the speed of the solitary wave V_s and the maximal force F_m. (Figure adapted from [9])

between beads is assumed to follow Hertz law, i.e., the contact force between beads k and $k-1$ is $F_{k-1,k} = \frac{D^{1/2}E}{3(1-\nu^2)}\delta_k^{3/2}$, where D, E, ν are the diameter, bulk elastic modulus, and Poisson ratio of the particles—this formula is easily generalized to heterogeneous chains [70]. In the above equation δ_k is given by $\delta_k \equiv \max\{u_{k-1} - u_k, 0\}$, where u_k is the displacement of the k-th particle from its equilibrium configuration. Then, the equations of motion that describe the system are

$$M_k \ddot{u}_k = \frac{D^{1/2}E}{3(1-\nu^2)}(\delta_k^{3/2} - \delta_{k+1}^{3/2}) \tag{5.1}$$

where M_k stands for the mass of the k-th particle. The application of this model to the experimental setup described above predicts the formation of solitary waves that travel along the chain with a speed V_s proportional to $F_m^{1/6}$, which is very close to the experimental value $V_s \propto F_m^{0.17}$. However, as shown in Figure 5.2, the experimentally-observed decay behavior of the force is not captured by this model. This discrepancy has recently motivated the inclusion of dissipative effects in the model, such as friction, plasticity, viscoelasticity, and viscous drag [9, 14, 31, 72, 73, 78]. Among these attempts of improving the predictability of the simulations, Daraio and co-workers [9] proposed for the first time a model able to capture both qualitatively and quantitatively the decay and wave shape observed experimentally. Their model includes dissipation in the form of a discrete Laplacian

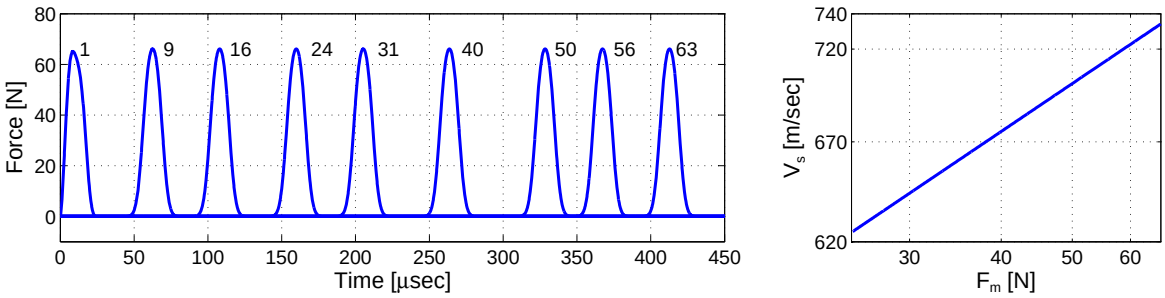


Figure 5.2: Set of numerical results for a one-dimensional chain composed of 70 steel beads with diameter $D = 0.00476$ m and impacted by a steel bead with $v_{\text{imp}} = 1.77$ m/s. The left side shows the time history of the forces predicted by the equations of motion (5.1). The right side shows the model estimation of solitary wave speed V_s versus maximal force F_m.

in the velocities with uniform exponent α and a material-dependent prefactor $\gamma < 0$, that is

$$M_k \ddot{u}_k = \frac{D^{1/2} E}{3(1-\nu^2)} (\delta_k^{3/2} - \delta_{k+1}^{3/2}) + M_k \gamma s |\dot{\delta}_k - \dot{\delta}_{k+1}|^{\alpha} \tag{5.2}$$

where $s = \text{sign}(\dot{\delta}_k - \dot{\delta}_{k+1})$, and the phenomenological parameters α and γ derive from best fitting with experimental observations. The dashed green curve in Figure 5.1 shows the decay behavior predicted by the model for $\alpha = 1.81$ and $\gamma = -5.58$, the best estimates for steel [9]. However, in order to successfully pursue the design of engineering devices that exploit the unique wave dynamics of granular crystals, a first-principles description that does not rely on empirical parameters but rather just on the knowledge of particle's geometry and material properties is required.

The work presented in this chapter is concerned with the formulation of first-principles predictive models of granular crystal dynamics. In particular, we investigate the role of vibrational-energy trapping effects in the overall dynamic behavior of one-dimensional chains under small to moderate impact velocities. To this end, we formulate two models whose only inputs are the geometry and the elastic material properties of the individual particles that comprise the granular crystal. The first model resolves the fine mesoscale structure of dynamic collisions and explicitly accounts for the vibrational kinetic energy retained in each bead as the solitary wave propagates along the chain, hence the *mesoscopic* nature of the approach. For small to moderate impact velocities, the model is conservative and the inclusion of dissipative effects is not required. We achieve these properties by

abandoning the particle mechanics approach and adopting a three-dimensional finite-element model. Specifically, our first approach is a dynamic contact problem of three-dimensional deformable elastic bodies that interact with one another over time.

The second approach proposed in this work is a one-dimensional regularized contact model where the vibrational energy that remains trapped after impact is subsumed under the concept of a coefficient of restitution. For small to moderate impact velocities, the variation of this coefficient with the impact velocity is a geometry and material dependent property that solely accounts for mesoscopic dynamic effects and that can be obtained from an experimental or numerical campaign of head-on collisions. We motivate this approach by a one-dimensional regularization of the three-dimensional contact problem and, in particular, we adopt the compliant normal-force model proposed by Hunt and Crossley [37]—where a damping coefficient is obtained from the coefficient of restitution. Thus, the one-dimensional regularized contact model is inherently energy-consistent and momentum-preserving, and it has the desirable properties of having as only inputs the geometry and the material properties of individual particles.

The chapter is organized as follows. In Section 5.2 we present our mesoscopic approach to granular crystal dynamics, we describe the three-dimensional finite-element model in Section 5.2.1 and the one-dimensional regularized contact model in Section 5.2.2. The verification and validation of these models are discussed in Section 5.3. For validation purposes, we use a set of experimental observations for a monodisperse one-dimensional chain of 28 stainless steel beads with diameter 0.00952 m (3/8") and impacted by a stainless steel bead with head-on velocity of magnitude 0.44271 m/s. Finally, a summary and concluding remarks are collected in Section 5.4.

5.2 Mesoscopic approach to granular crystal dynamics

The hypothesis and limitations of the commonly used particle mechanics approach to granular crystal dynamics are well-known but may stand a brief review as they serve as motivation to the mesoscopic approach presented in this section.

Firstly, the application of the static Hertz law to dynamic problems implies the following re-

strictions: (i) the maximum stress achieved in the vicinity of the contact must be less than the elastic limit, (ii) the size of the contact surface is much smaller than the radii of curvature of each particle, and (iii) the characteristic time of the problem is much smaller than the natural oscillation period of the elastic particle. However, these restrictions could be broken when, among other things, (i) particles are made of rubber-like materials, e. g., a neo-Hookean material model is required to describe the stress-strain behavior of polytetrafluoroethylene (PTFE) beads, (ii) the vibration produced by the collision is not negligible, and (iii) the relative motion between particles does not occur along the line joining centers of mass, e. g., bi-dimensional or three-dimensional arrangements of not necessarily spherical particles.

Secondly, dissipative effects are neglected in the equations of motion (5.1). Dissipative effects related to friction are negligible because highly-packed granular lattices are designed to favor axial stress propagation reducing then frictional and rotational dynamics. Dissipative effects during head-on collision have three main components: (i) the vibrational energy that remains *trapped* after impact, (ii) the energy loss due to plastic deformation or fracture, and (iii) the energy dissipation due to viscoelastic behavior of the material. A distinguishing characteristic between these three forms of energy loss is that the vibrational energy can be partially recovered by the particles as a result of subsequent collisions while local losses are permanent. The coefficient of restitution is then classically introduced to account for such losses and is generally treated as an empirical material-dependent constant. However, it is well-known that the coefficient of restitution is not only a function of the elastic properties but it also depends on the dimension of the particles and the colliding velocity.

The formulation of a first-principle predictive model of granular crystal dynamics which has as only inputs the geometry and the material properties of the individual particles that comprise the system is a challenging task in general. In this work, we specifically investigate the role of energy-trapping effects in one-dimensional granular crystals at low-impact velocity conditions. Under these conditions, local permanent energy losses are negligible and therefore a predictive model only has to account for the vibrational energy that remains trapped after impact, i.e., the model has to conserve

the total energy and momentum maps of the system. We achieve these properties by abandoning the particle mechanics approach and adopting a three-dimensional finite-element model. Specifically, the model corresponds to a dynamic contact problem of three-dimensional deformable elastic bodies that interact with one another over time.

In the same spirit, we formulate a one-dimensional regularized contact model where the vibrational energy that remains trapped after impact is subsumed under the concept of a coefficient of restitution. Despite being attractive for its simplicity, this approach is inherently energy-consistent and momentum-preserving, and has the desirable properties of having as only inputs the geometry and the material properties of individual particles.

These two models are described next in turn.

5.2.1 Three-dimensional finite-element model

The first model presented in this work makes use of a three-dimensional finite-element mesh of the one-dimensional chain of beads. The mesh of each bead in the system comprises $27,200$ 4-node tetrahedral isoparametric elements and $5,559$ nodes, Figure 5.3. The contact constraint between beads is enforced using a penalty energy function as shown in Gonzalez *et. al* [24] and in Chapter 2. The elastodynamic problem is then described by a finite-dimensional Lagrangian system of the form

$$L(q,\dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - V(q) - I_{\mathcal{C}}(q) \tag{5.3}$$

where the generalized coordinates q are the coordinates of the nodes in the deformed configuration, M is the mass matrix, $V(q)$ is the elastic energy, and $I_{\mathcal{C}}(q)$ is the indicator function of the set of admissible configurations $\mathcal{C}$, i.e., the impenetrability or contact constraint. The striker is imparted a uniform initial velocity which is referred to as impact velocity v_{imp}. The elastic behavior of the beads is described by a strain-energy density of the form

$$W(F) = \frac{\lambda_0}{2}(\log J)^2 - \mu_0 \log J + \frac{\mu_0}{2}\operatorname{tr}\left(F^T F\right) \tag{5.4}$$

which describes a neo-Hookean solid extended to the compressible range, and where λ_0 and μ_0 are the Lamé constants. The Lagrangian system (5.3) is then momentum and energy preserving. Finally, the trajectories of the strongly nonlinear dynamical system are obtained by numerical time integration.

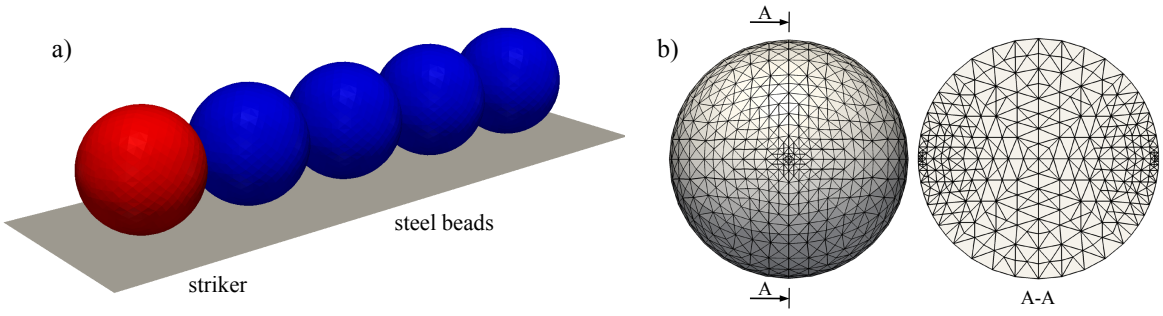


Figure 5.3: One-dimensional chain of beads. a) Three-dimensional finite-element model. b) Detail of the finite-element mesh of one spherical bead.

Strongly nonlinear three-dimensional systems, with complex dynamic contact and multiple time scales that coexist in time and space, pose great challenges for numerical time integrators. We address these challenges with asynchronous energy-stepping integrators. These time integrators, developed by Gonzalez *et al.* in [26] and presented in Chapter 4, are energy-momentum conserving, symplectic, and convergent with automatic and asynchronous selection of time step in each subdomain—each bead and each contact constraint is a partition of the domain. The scheme conserves all the exact momentum maps of the Lagrangian system and accounts for the exact momentum fluxes across subdomain boundaries. Furthermore, asynchronous energy-stepping is effective and efficient in solving the dynamic behavior of a chain of elastic beads, as shown in Chapter 4.

Of particular interest to granular crystal dynamics are conservation of total energy and linear momentum. The total energy of the system is the sum of the kinetic energy, $\mathrm{KE}(\dot{q}) = \frac{1}{2}\dot{q}^T M\dot{q}$, and the potential energy $V(q)$. The total linear momentum is given by $J = M\dot{q}$. It is interesting to observe that the kinetic energy and the linear momentum of the system can also be expressed as the contribution of all individual particles k in the one-dimensional granular crystal, i.e., KE_k and J_k,

given by

$$\mathrm{KE}_k = \sum_{q_i \in \mathcal{V}_k} \frac{1}{2} m_i \|\dot{q}_i\|^2 \tag{5.5}$$

$$J_k = \sum_{q_i \in \mathcal{V}_k} m_i \dot{q}_i = M_k \langle \dot{q} \rangle_k \tag{5.6}$$

where M_k, $\langle \dot{q} \rangle_k$, and $\mathcal{V}_k$ are the total mass, the mean velocity and the volume of particle k, and m_i is the mass of node i. We now define the vibrational kinetic energy of each bead as

$$\mathrm{VKE}_k = \sum_{q_i \in \mathcal{V}_k} \frac{1}{2} m_i \|\dot{q}_i - \langle \dot{q} \rangle_k\|^2 = \mathrm{KE}_k - \langle \mathrm{KE} \rangle_k \tag{5.7}$$

where $\langle \mathrm{KE} \rangle_k = \frac{1}{2} M_k \|\langle \dot{q} \rangle_k\|^2$ is the kinetic energy associated with the rigid motion of the particle. The additive decomposition of the kinetic energy into vibrational and rigid-motion components will allow for investigating the role of energy-trapping effects in granular crystals at low-impact velocity conditions.

For the purpose of validating the model, it is essential to identify and compute forces that resemble those measured by embedded piezosensors in experimental setups. Daraio and co-workers [15] measure an average value of forces F_k^{left} and F_k^{right} acting on the contact points of the particle, that is

$$F_k^{av} = \frac{\left| F_k^{left} \cdot n \right| + \left| F_k^{right} \cdot n \right|}{2} = \frac{1}{2} \left(F_k^{left} - F_k^{right} \right) \cdot n \tag{5.8}$$

where n is a unit vector aligned with the chain of beads, Figure 5.4. This averaged force is related with the external forces acting on the beads by the following relationship

$$F_k^{av} = \frac{1}{2} F_k^{ext} \cdot n + \sum_{j=k+1}^{N_{\mathrm{beads}}} F_j^{ext} \cdot n \tag{5.9}$$

where N_{beads} is the number of beads in the chain and F_k^{ext} is the external force acting on particle k, i.e.,

$$F_k^{ext} = F_k^{left} + F_k^{right} = \frac{d}{dt} J_k = M_k \frac{d}{dt} \langle \dot{q} \rangle_k \tag{5.10}$$

Relationship (5.9) is only valid for granular crystals that do not interact with walls, i.e., for systems

that preserve the total linear momentum. However, it can be readily extended to consider other cases.

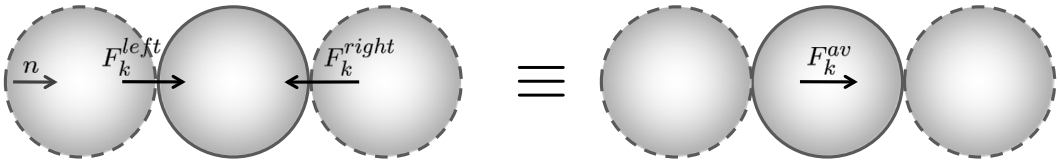


Figure 5.4: Measured forces acting on a bead.

5.2.2 One-dimensional regularized contact model

The second model presented in this work is motivated by a one-dimensional regularization of the three-dimensional contact problem presented in the previous section. We first recall that Hamilton's principle states that trajectories $q(t)$ of the Lagrangian (5.3) are the stationary points of the action integral of the system, that is

$$0 = \delta \int L(q, \dot{q})\ dt = \delta \int \left[\langle \mathrm{KE} \rangle(\dot{q}) + \mathrm{VKE}(\dot{q}) - V(q) - I_{\mathcal{C}}(q)\right]\ dt \tag{5.11}$$

Then, for a one-dimensional chain of particles, we define the displacement $u(t)$ and the velocity $\dot{u}(t)$ as

$$u(t) := \langle q \rangle|_t \cdot n - \langle q \rangle|_0 \cdot n \tag{5.12}$$

$$\dot{u}(t) := \langle \dot{q} \rangle|_t \cdot n \tag{5.13}$$

and we assume other components of the displacement are zero, that is $\langle q \rangle = (\langle q \rangle \cdot n)\ n$. This intuitive dimensional reduction suggests recasting the three-dimensional Lagrangian system into a one-dimensional mechanical system with forcing. Thus, trajectories $u(t)$ are given be the Lagrange-d'Alembert principle, i.e.,

$$0 = \delta \int L(u, \dot{u})\ dt \quad \rightarrow \quad 0 = \delta \int \left[\langle \mathrm{KE} \rangle(\dot{u}) - \bar{V}(u)\right]\ dt + \int F\,(u, \dot{u}) \cdot \delta u\ dt \tag{5.14}$$

where $F(u, \dot{u})$ is the forcing term and $\bar{V}$ is a one-dimensional regularized potential. The resulting equations of motion are then given by

$$M_k \ddot{u}_k = -\nabla_k \bar{V}(u) + F_k(u, \dot{u}) \tag{5.15}$$

where forces are approximated by a regularized contact model or compliant contact-force model. It is worth noting that in our approach the forcing term results in an effective dissipation of the vibrational energy retained in the particles during and after the collision.

For direct central and frictionless impacts of two particles, Hunt and Crossley [37] proposed a compliant normal-force model of the form

$$m(\ddot{u}_1 - \ddot{u}_2) = -\kappa\delta^n - (\alpha\kappa\delta^n)\dot{\delta} \tag{5.16}$$

where $\delta = \max\{u_1 - u_2, 0\}$ and $\dot{\delta} = \dot{u}_1 - \dot{u}_2$ are the penetration depth and speed, α is the damping factor, κ is the spring constant, and m is the effective mass (i.e., $m^{-1} = M_1^{-1} + M_2^{-1}$). This regularized contact model falls squarely within the dimensional reduction proposed above and the equations of motion (5.15). Indeed, the potential energy is $\bar{V}(\delta) = \frac{\kappa}{n+1}\delta^{n+1}$ and the dissipative forcing term is $F(\delta, \dot{\delta}) = -(\alpha\kappa\delta^n)\dot{\delta}$. It is worth noting that the potential energy is chosen in analogy with Hertz's theory, which is a good regularized model for the contact force problem of hard elastic bodies where the contact region remains small and no vibrational modes are exited during collision. Then, for heterogeneous pairs of linearly elastic spheres, $n = 3/2$ and the spring constant κ is well-known (cf., e. g., [70]). The damping factor α is generally chosen to ensure that the energy dissipated during impact is consistent with the energy loss subsumed in the coefficient of restitution. Additionally, in order to prevent the contact model from applying tensile force, α must verify

$$\delta^n + \alpha\delta^n\dot{\delta} \geq 0 \quad , \quad \forall\ \dot{\delta} \tag{5.17}$$

Many researchers have proposed approximate [37, 48, 56] and exact [23] relationships between α

and the coefficient of restitution (see, for example, [82] for a detailed comparison of the predictions of these models with low speed impact measurements). In this work we adopt the exact solution proposed by Gonthier and co-workers [23]. To this end we integrate the equation of motion (5.16) as follows

$$\int_{v_\mathrm{i}}^{v_\mathrm{o}} \frac{\ddot{\delta}}{1+\alpha\dot{\delta}}\, d\dot{\delta} + \frac{\kappa}{m}\int_{x_\mathrm{i}}^{x_\mathrm{o}} \delta^n\, d\delta = 0 \tag{5.18}$$

where x_i, x_o, v_i and v_o are the penetration depths and speeds at the start and the end of the collision (i.e., $x_\mathrm{i} = x_\mathrm{o} = 0$). We also define the effective coefficient of restitution $e = -\frac{v_\mathrm{o}}{v_\mathrm{i}}$ that allows for expressing the exact solution of (5.18) as an implicit relation for α, given v_i and e, that is

$$\frac{1+\alpha v_\mathrm{i}}{1-\alpha v_\mathrm{i} e} = \mathrm{e}^{\alpha v_\mathrm{i}(1+e)} \tag{5.19}$$

The solution for α has multiple branches, however a valid solution should satisfy inequality (5.17).

The model proposed by Gonthier and co-workers, i.e., (5.19), is momentum-preserving and energy-consistent for any given value of the coefficient of restitution. Therefore, the model can be used to simulate fully elastic to nearly-plastic impacts. At low impact velocities and for most linear elastic materials, experimental observations [22] traditionally show that the coefficient of restitution can be approximated by $e_\mathrm{emp} = 1 - c_1 v_\mathrm{i}$, where c_1 is an empirical coefficient. However, recent experimental data obtained by Zhang and Sharf [82] suggest a nonlinear fit of the form $e_\mathrm{emp} = 1 - c_1 v_\mathrm{i}^{c_2}$ provides a better approximation. Then, the damping factor is only a function of the speed at the start of the collision, i.e., $\alpha = \alpha(v_\mathrm{i})$, and the empirical coefficients c_1, c_2.

The application of Hunt-Crossley's regularized contact model to equations of motion (5.15) results in

$$M_k \ddot{u}_k = \frac{D^{1/2}E}{3(1-\nu^2)}\left[\delta_k^{3/2}(1+\alpha_k\ \dot{\delta}_k) - \delta_{k+1}^{3/2}(1+\alpha_{k+1}\ \dot{\delta}_{k+1})\right] \tag{5.20}$$

where $\alpha_k = \alpha_k(v_{\mathrm{i},k})$ is given by Gonthier's energy-consistent model. The function $e_\mathrm{emp}(v_\mathrm{i})$ can be obtained from an experimental or numerical campaign of head-on collisions over a range of initial impacting velocities. The value for $v_{\mathrm{i},k}$ is approximated by the largest attained relative velocity

between $(k-1)$-th and k-th particles (see Appendix 5.4 for evolutionary equations for v_{i}). It bears emphasis that e_{emp} accounts for the vibrational energy retained in the bead after the head-on collision, and that the equations of motion (5.20) reduce to (5.1) when $e_{\mathrm{emp}} = 1$ for all v_{i}.

Finally, the one-dimensional model is integrated over time with variational integrators which accurately capture the energy behavior of forced mechanical systems [42].

5.3 Verification and validation

The verification of the models presented in the previous section is twofold: (i) the assessment of the convergence and accuracy of the numerical solutions to the exact solutions of the models, and (ii) the assessment of the accuracy of the one-dimensional regularized contact model as an approximation of the three-dimensional finite-element model. Our verification and validation strategy is depicted in Figure 5.5.

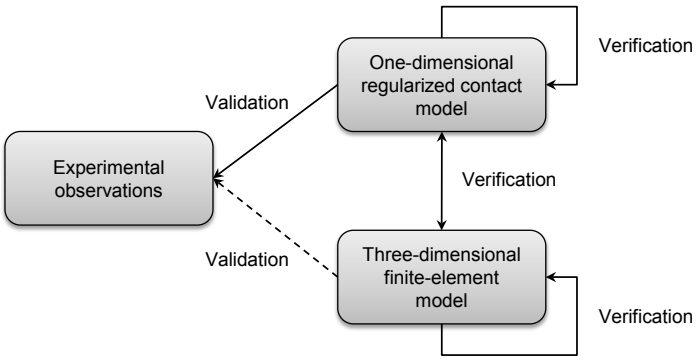


Figure 5.5: Verification and validation strategy.

For validation purposes, we use a set of experimental observations for a monodisperse one-dimensional chain of 28 stainless steel beads with diameter $D = 0.00952$ m (3/8") and impacted by a stainless steel bead with $v_{\mathrm{imp}} = 0.44271$ m/s. The time history of forces measured by 15 calibrated piezosensors is presented in Figure 5.6—piezosensors are placed in beads 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 27, 28, where the striker is referred as the first bead. Characteristically, the measured force decays as the solitary wave propagates along the chain. The material properties of the beads are $E = 213$ GPa, $\nu = 0.30$, and the density is $\rho = 7780.0$ kg/m^3.

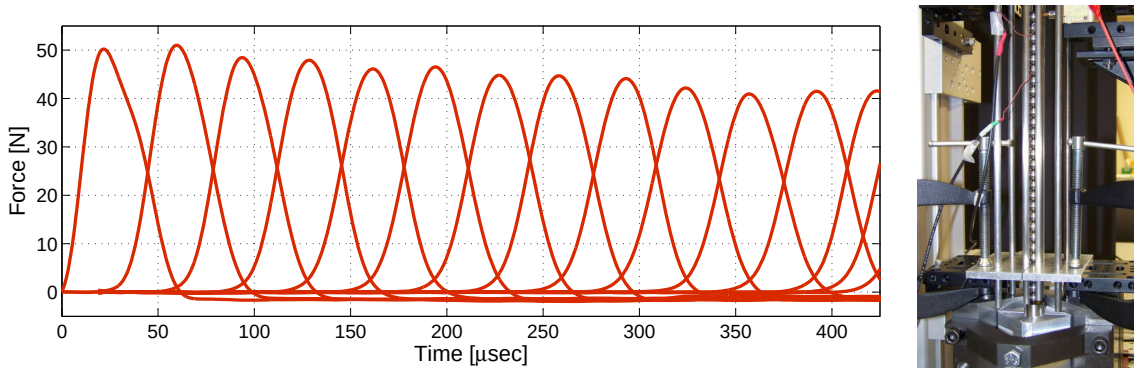


Figure 5.6: Left: Time history of the forces measured by 15 sensors for a one-dimensional chain composed of 28 stainless steel beads with diameter $D = 0.00952$ m and impacted by a stainless steel bead with $v_{\text{imp}} = 0.44271$ m/s. Right: Experimental setup. Courtesy of Dr. Jinkyu Yang and Prof. Chiara Daraio.

5.3.1 Verification

Convergence, accuracy, long-term behavior and conservation properties of energy-stepping and asynchronous energy-stepping are carefully studied in Gonzalez *et al.* [24,26], and Chapters 2 and 4. The numerical behavior of variational integrators for forced mechanical systems is analyzed in Kane *et al.* [42].

We assess the accuracy of the one-dimensional regularized contact model (Section 5.2.2) as an approximation of the three-dimensional finite-element model (Section 5.2.1) by means of numerical experimentation. Specifically, we study the head-on impact of two stainless steel beads with diameter $D = 0.00952$ m and impact velocity $v_{\text{imp}} = 0.44271$ m/s. The three-dimensional model uses the finite-element mesh shown in Figure 5.3 and it is integrated in time with asynchronous energy-stepping integrators and an energy step of magnitude $h = 2 \cdot 10^{-8}$ J. The model predicts a coefficient of restitution of $e = 0.975$ which is in turn used as an input of the one-dimensional regularized model.

The regularized contact model is designed to exactly predict the coefficient of restitution as a consequence of solving Gonthier's model (5.19) for computing the damping factor α. However, an accurate prediction of the dynamic behavior is required to successfully apply the model to granular crystal dynamics. To this end, we compare the three-dimensional and the one-dimensional models in their predictions of the evolution of: (i) incoming and target beads linear momenta, (ii) kinetic,

potential and total energies [1], and (iii) the external or contact force. Figure 5.7 compares the time histories of such three aspects of the dynamic behavior of the system. The good agreement between the evolution in time of the total energy of the one-dimensional model and the total energy of the three-dimensional model *adjusted* by the vibrational energy retained in the system is noteworthy. It also bears emphasis the good agreement in the force amplitude and contact duration. A small discrepancy is observed at times where the potential energy is small, i.e., at small deformations, and this discrepancy could be attributed to the lack of refinement in the finite-element mesh for such conditions. Figure 5.7 additionally shows the exact conservation of linear momentum in the both models (grey curves), and the exact conservation of total energy in the three-dimensional model (grey curve). We again emphasize that the vibrational energy that remains trapped after impact is subsumed, in the one-dimensional model, under the concept of a coefficient of restitution.

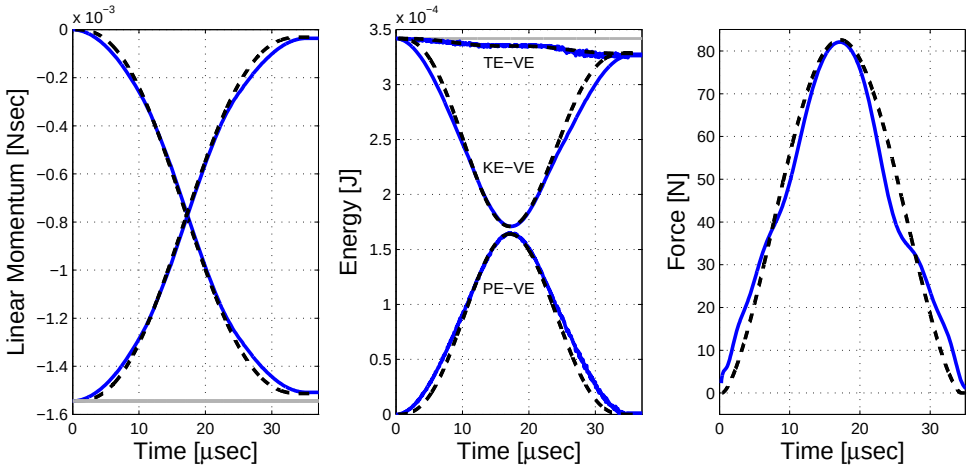


Figure 5.7: Head-on impact of two stainless steel beads with diameter $D = 0.00952$ m and impact velocity $v_{\text{imp}} = 0.44271$ m/s. Predictions of the three-dimensional finite-element model (dashed black curves) and the one-dimensional regularized contact model (blue curves).

In order to assess the accuracy of these models in the context of granular crystal dynamics, we now study a one-dimensional chain of stainless steel beads with diameter $D = 0.00952$ m and impact velocity $v_{\text{imp}} = 0.44271$ m/s. Figure 5.8 shows the solutions of the one-dimensional regularized contact model and the three-dimensional finite-element model—the averaged force is computed as described in Section 5.2.1. The good agreement of the models is evident in the figure, although

[1] The vibrational energy of the three-dimensional model is explicitly computed by (5.7) and subtracted from the system in order to compare the predictions with the one-dimensional regularized model.

the discrepancies observed at small deformations play a more significant role due to the higher concurrency of this dynamic problem. Nonetheless, the accord between the evolution in time of the total energy of the one-dimensional model and the total energy of the three-dimensional model *adjusted* by the vibrational energy retained in the system is noteworthy.

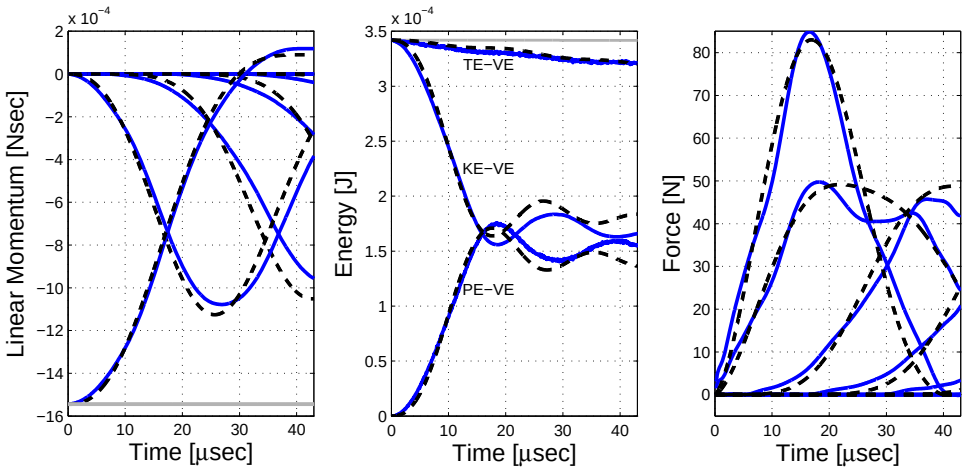


Figure 5.8: Chain of stainless steel beads with diameter $D = 0.00952$ m, impacted by a stainless steel bead with $v_{\text{imp}} = 0.44271$ m/s. Predictions of the three-dimensional finite-element model (dashed black curves) and the one-dimensional regularized contact model (blue curves).

5.3.2 Validation

We now proceed to validate the one-dimensional regularized contact model (Section 5.2.2) with the experimental observations described above and shown in Figure 5.6. In order to simulate the dynamic behavior of the one-dimensional chain of beads, we additionally require the coefficient of restitution $e(v_{\text{i}})$ for the specific geometry and the material properties of the particles that comprise the granular crystal. Zhang *et al.* presented in [82] a set of experimentally obtained coefficients of restitution for low speed impacts of a chrome-steel ball with diameter 0.0508 m ($E = 210$ GPa, $\nu = 0.30$) on a steel cylinder with diameter 0.0508 m ($E = 213$ GPa, $\nu = 0.29$). Figure 5.9 shows such a set of experimental results together with a nonlinear fit of the form $e_{\text{emp}} = 1 - 0.0626\ v_{\text{i}}^{0.3860}$. In this work we perform a numerical campaign of head-on impacts with the three-dimensional finite-element model presented in Section 5.2.1. These numerical results are also shown in Figure 5.9 together with a nonlinear fit of the form $e_{\text{emp}} = 1 - 0.0370\ v_{\text{i}}^{0.4101}$. As expected for similar material

properties, the coefficient of restitution is larger for a smaller diameter. It is worth noting that both models presented in this work are intertwined in the validation, as it was suggested in Figure 5.5.

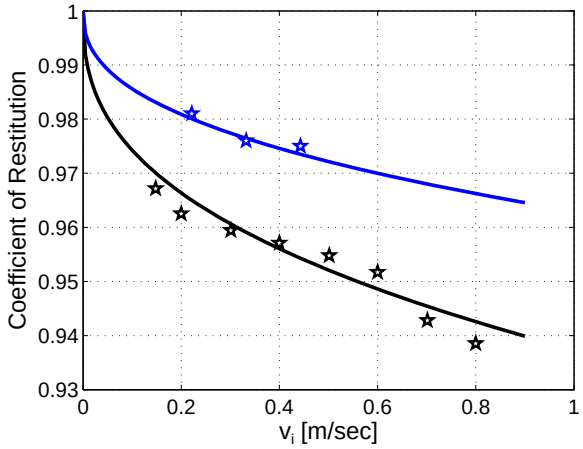


Figure 5.9: Coefficient of restitution $e(v_{\mathrm{i}})$. Black stars are experimental observations by Zhang [82], the black curve is the best fit $e = 1-0.0626\ v_{\mathrm{i}}^{0.3860}$. Blue stars correspond to the numerical campaign of head-on impacts of stainless steel beads with diameter $D = 0.00952$ m, the blue curve corresponds to $e = 1 - 0.0370\ v_{\mathrm{i}}^{0.4101}$.

The time history of the forces predicted by the one-dimensional regularized contact model—with equations of motion given by (5.20)—and the forces predicted by the traditional particle mechanics approach—with equations of motion given by (5.1)—are shown in Figure 5.10. It is evident in the figure that the traditional particle mechanics approach does not predict the characteristic decay of the force as the solitary wave propagates along the chain. In sharp contrast, the mesoscopic approach proposed in this work successfully captures such behavior. Furthermore, the agreement between the numerical results and the experimental observations is remarkable, as shown in Figure 5.11.

The one-dimensional regularized contact model is momentum-preserving and energy-consistent. The first property is observed in Figure 5.12, namely the total linear momentum of the system is a constant of motion though the shape of the individual contributions J_k changes as the solitary wave propagates along the chain. In contrast, the traditional particle mechanics approach predicts the same profile for all J_k, after a short transient period. The second property is observed in Figure 5.13, and it is worth noting that the decaying behavior in the total energy of the system resembles the vibrational energy that remains trapped in the beads.

Finally, the one-dimensional regularized contact model predicts all the well-known qualitative

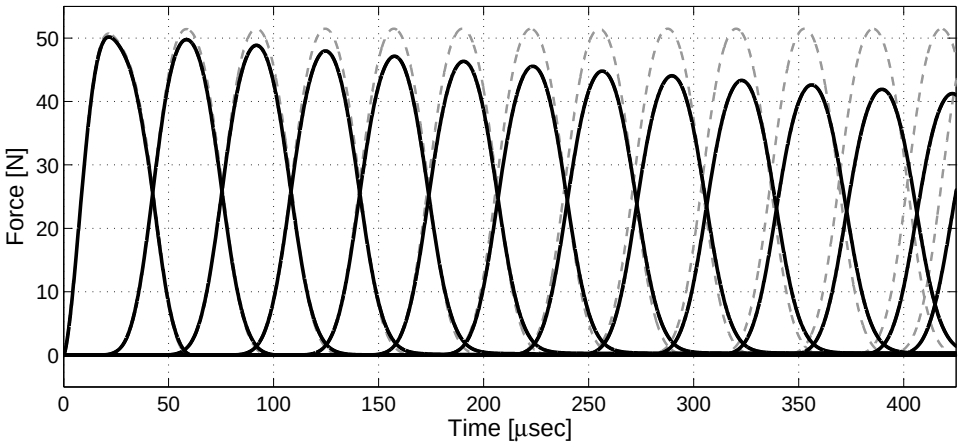


Figure 5.10: One-dimensional regularized contact model of a one-dimensional chain composed of 28 stainless steel beads with diameter $D = 0.00952$ m and impacted by a stainless steel bead with $v_{\text{imp}} = 0.44271$ m/s. Time history of the forces predicted by the model—dashed grey curves correspond to the predictions of equations of motion (5.1).

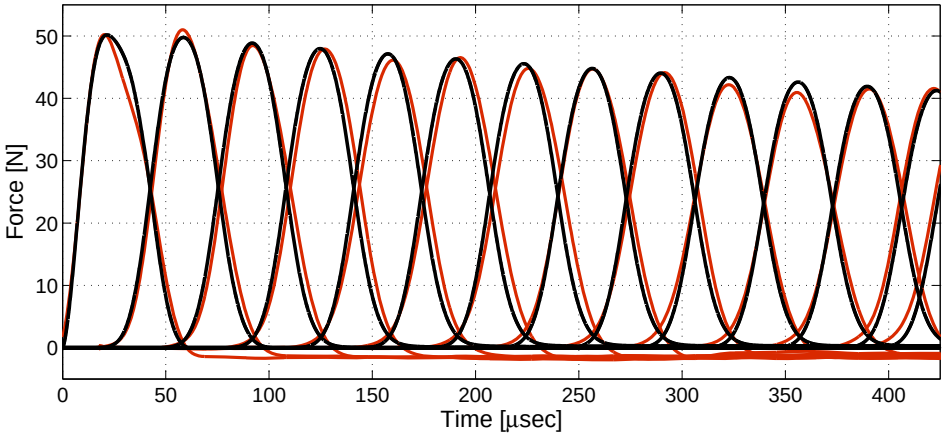


Figure 5.11: One-dimensional regularized contact model of a one-dimensional chain composed of 28 stainless steel beads with diameter $D = 0.00952$ m and impacted by a stainless steel bead with $v_{\text{imp}} = 0.44271$ m/s. Time history of measured forces (red curves) and numerical predictions (black curves).

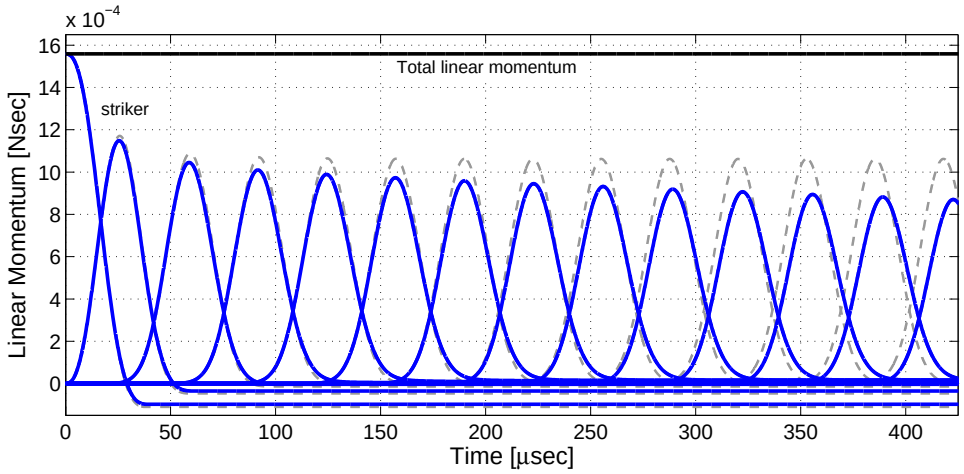


Figure 5.12: One-dimensional regularized contact model. Conservation of total linear momentum and time history of individual J_k—dashed grey curves correspond to the predictions of equations of motion (5.1).

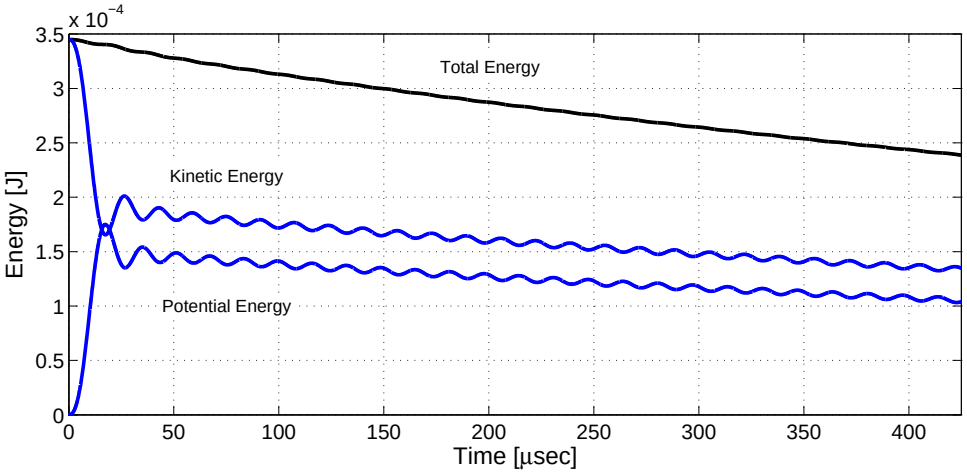


Figure 5.13: One-dimensional regularized contact model. Time history of total, kinetic and potential energies.

behavior of one-dimensional granular crystal dynamics. In particular, we observe that:

- The solitary wave width remains constant after a short transient period (see Figure 5.14). The width is measured by the full width at half maximum and it is in agreement with experimental and numerical observations that report a value of about 2.1 particles [65, 70]. Thus, the mesoscopic approach proposed in this work predicts solitary waves with a finite width that is independent of the solitary wave amplitude.

- The decay of the force follows $F_m \propto \mathrm{e}^{-\eta k}$ (see Figure 5.15). This behavior is in agreement with experimental results reported elsewhere (see, for example, [9]).

- The wave speed is given by $V_s \propto F_m^{0.157}$ (see Figure 5.16). This is in good agreement with theoretical predictions [64, 65] and experimental data [9, 15, 16].

5.4 Summary and discussion

We have presented a mesoscopic approach to granular crystal dynamics, which comprises a three-dimensional finite-element model and a one-dimensional regularized contact model. The approach aims at investigating the role of vibrational-energy trapping effects in the overall dynamic behavior of one-dimensional chains under small to moderate impact velocities. The three-dimensional finite-element model resolves the fine mesoscale structure of dynamic collisions and explicitly accounts for the vibrational kinetic energy retained in each bead as the solitary wave propagates along the chain. The resulting strongly nonlinear three-dimensional system is conservative (i.e., the total energy of the system is a constant of motion), and it poses great challenges for numerical time integrators. We have successfully coped with these challenges by using asynchronous energy-stepping integrators. The one-dimensional regularized contact model accounts for mesoscopic dynamic effects by means of a restitution coefficient, i.e., the vibrational energy that remains trapped after impact is subsumed under the concept of a coefficient of restitution. We have specifically adopted the compliant normal-force model proposed by Hunt and Crossley, and the damping-factor model proposed by Gonthier and co-workers. The resulting one-dimensional model is momentum-preserving and energy-consistent,

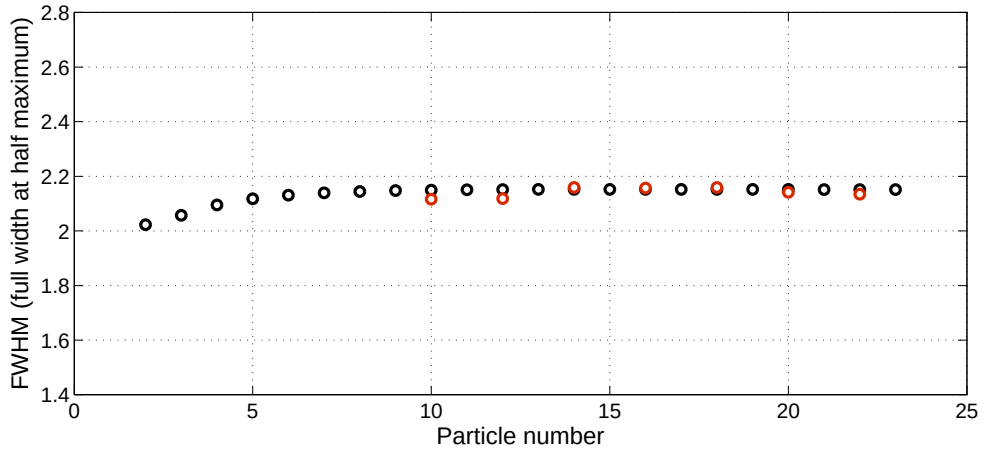


Figure 5.14: One-dimensional model. Evolution of solitary wave width (full width at half maximum) as a function of bead number—measured values in red.

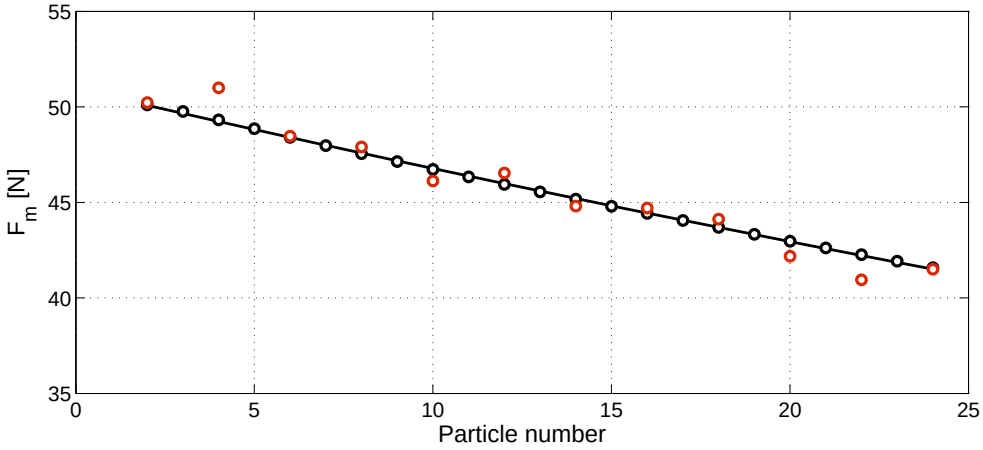


Figure 5.15: One-dimensional model. Evolution of the maximum force F_m as a function of the bead number k—measured values in red. The curve corresponds to the best fit $F_m = 50.947\ \mathrm{e}^{-0.0085411k}$.

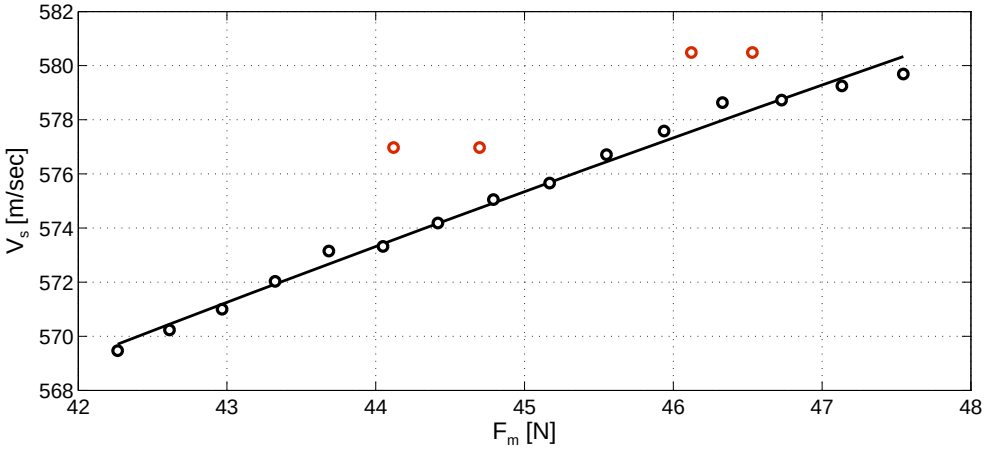


Figure 5.16: One-dimensional. Speed of the solitary wave V_s versus the maximum force F_m—measured values in red. The curve corresponds to the best fit $V_s = 316.52\ F_m^{0.15698}$.

and it is integrated over time with variational integrators which accurately capture the energy behavior of forced mechanical systems. The only inputs of these models are the geometry and the elastic material properties of the individual particles that form the granular crystal (e. g., we have obtained the variation of the coefficient of restitution with the impact velocity from a numerical campaign of head-on collisions).

We have also presented a detailed verification and validation of the mesoscopic approach, which include: (i) the assessment of the convergence and accuracy of the numerical solutions to the exact solutions of the models, (ii) the assessment of the accuracy of the one-dimensional regularized contact model as an approximation of the three-dimensional finite-element model, (iii) the one-to-one comparison of the experimental and simulated time histories of averaged forces in a one-dimensional chain of 28 stainless steel beads with diameter 0.00952 m and a small impact velocity of 0.44271 m/s. The good agreement of the latter and the ability of the model to predict well-known properties of one-dimensional granular crystal dynamics (e. g., the formation of solitary waves with a finite width that is independent of the solitary wave amplitude, the exponential decay of the force as the solitary wave propagates along the chain, and the $F_m^{1/6}$ scaling of the solitary wave speed) are remarkable.

We close by pointing out some limitations of our analysis and possible avenues for extensions of the approach.

Firstly, our mesoscopic approach only accounts for vibrational-energy trapping effects and it therefore relies on the assumption of negligible permanent energy loses in the system. However, the strain-energy density employed in the finite-element model can be easily extended to account for elasto-viscoplastic materials, e. g., steel and teflon at high-impact velocities, or even to account for brittle materials, e. g., glass. Daraio and co-workers have reported a considerable amount of experimental observations that provide a wealth of data for validating a model under such conditions, though beyond the scope of this work.

Secondly, the application of the three-dimensional finite-element model to heterogeneous systems composed of not necessarily spherical particles, and arranged in bi-dimensional or three-dimensional configurations, is a worthwhile direction of future research.

Appendix: Loading/unloading conditions for $v_{\rm i}$

The damping factor $\alpha(v_{\rm i})$ in the equation of motion

$$m(\ddot{u}_1 - \ddot{u}_2) = -\kappa\delta^n \left[1 + \alpha(v_{\rm i})\ \dot{\delta}\right] \tag{5.21}$$

is determined from

$$\frac{1 + \alpha v_{\rm i}}{1 - \alpha v_{\rm i}\ e(v_{\rm i})} = {\rm e}^{\alpha v_{\rm i}[1+e(v_{\rm i})])} \tag{5.22}$$

given the value of the impact speed $v_{\rm i} \geq 0$. Whether $v_{\rm i}$ is positive or zero depends on further conditions involving $(\delta, \dot{\delta})$, which we refer to as loading/unloading and touching/detaching conditions for $v_{\rm i}$. We present below these conditions for functions $v_{\rm i} : SBV([0,T]) \to \mathbb{R}$ in the space of Special Bounded Variation functions and for derivatives $a_{\rm i} := v_{\rm i}'$ that exist almost everywhere.

1. Loading/unloading conditions ($\delta > 0$).

   The irreversible nature of the percussion is captured by Kuhn-Tucker conditions

   $$a_{\rm i} \geq 0 \quad , \qquad \dot{\delta} - v_{\rm i} \leq 0 \quad , \qquad a_{\rm i}(\dot{\delta} - v_{\rm i}) = 0$$

   together with the consistency condition, if $\dot{\delta} - v_{\rm i} = 0$,

   $$a_{\rm i} = \ddot{\delta} \quad \text{if} \quad \ddot{\delta} \geq 0 \quad , \qquad a_{\rm i} = 0 \quad \text{if} \quad \ddot{\delta} < 0$$

2. Touching/detaching conditions ($\delta = 0$).

   Before and after the percussion of two beads, the value of $v_{\rm i}$ may experience a jump given by:

   - at the first instant of contact ($\dot{\delta} > 0$): $v_{\rm i} = \dot{\delta}$

   - after the collision ($\dot{\delta} \leq 0$): $v_{\rm i} = 0$

Bibliography

[1] VI Arnold, "Mathematical methods of classical mechanics", Second Edition. *Graduate Texts in Mathematics* **60**, Springer-Verlag, New York, 1980.

[2] R Abraham and JE Marsden, "Foundations of mechanics", Second Edition. Addison-Wesley Publishing Company Inc., Redwood City, CA., 1987.

[3] K-J Bathe, "Finite Element Procedures". Prentice-Hall, Englewood Cliffs, NJ, 1996.

[4] T Belytschko and D Schoeberle, "On the unconditional stability of an implicit algorithm for nonlinear structural dynamics", *Journal of Applied Mechanics*, **42**, 865–869, 1975.

[5] JJ Biesiadecki and RD Skeel, "Dangers of multiple time step methods", *Journal of Computational Physics*, **109**, 318–328, 1993.

[6] A Bliss, and FE Su, "Lower bounds for simplicial covers and triangulations of cubes", *Discrete Comput. Geom.*, **33**, 669–686, 2005.

[7] J Brandts, S Korotov, M Křížek, and J Šolc, "On nonobtuse simplicial partitions", *SIAM Review*, **51**:2, 317–335, 2009.

[8] MP Calvo and JM Sanz-Serna, "The development of variable-step symplectic integrators, with applications to the two-body problem", *SIAM J. Sci. Comput.*, **4**, 936–952, 1993.

[9] R Carretero-González, D Khatri, MA Porter, PG Kevrekidis, and C Daraio, "Dissipative solitary waves in granular crystals", *Physical Review Letters*, **102**, 024102, 2009.

[10] NJ Carpenter, RL Taylor, and MG Katona, "Lagrange constraints for transient finite-element surface-contact", *Int. J. Numer. Meth. Eng.*, **32**:1, 103–128, 1991.

[11] M-J Chien and ES Kuh, "Solving nonlinear resistive networks using piecewise-linear analysis and simplicial subdivision", *IEEE Transactions on Circuits and Systems*, **24**:6, 305-317, 1977.

[12] M Chien, "Searching for multiple solutions of nonlinear systems", *IEEE Transactions on Circuits and Systems*, **22**:10, 817–827, 1979.

[13] LO Chua, "Efficient computer algorithms for piecewise-linear analysis of resistive nonlinear networks", *IEEE Transactions on Circuits Theory*, **18**:1, 73–85, 1971.

[14] C Coste, E Falcon, and S Fauve, "Solitary waves in a chain of beads under Hertz contact", *Physical Review E*, **56**, 6104–6117, 1997.

[15] C Daraio, VF Nesterenko, EB Herbold, and S Jin, "Strongly nonlinear waves in a chain of Teflon beads", *Physical Review E*, **72**, 016603, 2005.

[16] C Daraio, VF Nesterenko, EB Herbold, and S Jin, "Tunability of solitary wave properties in one-dimensional strongly nonlinear phononic crystals", *Physical Review E*, **73**, 026610, 2006.

[17] C Daraio, VF Nesterenko, EB Herbold, and S Jin, "Energy trapping and shock disintegration in a composite granular medium", *Physical Review Letters*, **96**, 058002, 2006.

[18] T Fujisawa and ES Kuh, "Piecewise-linear theory of nonlinear networks", *SIAM J. Appl. Math.*, **22**:2, 307–328, 1972

[19] Z Ge and JE Marsden, "Lie–Poisson integrators and Lie–Poisson Hamilton–Jacobi theory", *Physics Letters A*, **133**:3, 134–139, 1988.

[20] B Gladman, M Duncan, and J Candy, "Symplectic integrators for long-term integrations in celestial mechanics", *Celestial Mech. Dynamical Astronomy*, **52**, 221–240, 1991.

[21] V Gol'dshtein and S Dubrovskiy, "Lemma Poincaré for $L_{\infty,loc}$-forms", 2007. arXiv:0712.1682v1 (http://arxiv.org/abs/0712.1682).

[22] W Goldsmith, "Impact: The theory and physical behaviour of colliding solids". Edward Arnold Ltd., London, 1960.

[23] Y Gonthier, J McPhee, C Lange, and JC Piedboeuf, "A regularized contact model with asymmetric damping and dwell-time dependent friction", *Multibody System Dynamics*, **11**, 209-233, 2004.

[24] M Gonzalez, B Schmidt, and M Ortiz, "Energy-stepping integrators in Lagrangian mechanics", *Int. J. Numer. Meth. Eng.*, DOI: 10.1002/nme.2753, 2009.

[25] M Gonzalez, B Schmidt, and M Ortiz, "Force-stepping integrators in Lagrangian mechanics", *Int. J. Numer. Meth. Eng.*, DOI: 10.1002/nme.2942, 2010.

[26] M Gonzalez, B Schmidt, and M Ortiz, "Asynchronous energy-stepping integrators in Lagrangian mechanics", in preparation, 2010.

[27] M Gonzalez, J Yang, C Daraio, and M Ortiz, "Mesoscopic approach to granular crystal dynamics", in preparation, 2010.

[28] WW Hager, "Updating the inverse of a matrix", *SIAM Review*, **31**:2, 221–239, 1989.

[29] E Hairer and C Lubich, "The life-span of backward error analysis for numerical integrators", *Numerische Mathematik*, **76**, 441–462, 1997.

[30] E Hairer, C Lubich, and G Wanner, "Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations", Second Edition. *Springer Series in Computational Mathematics* **31**, Springer-Verlag, Berlin, 2006.

[31] EB Herbold and VF Nesterenko, "Shock wave structure in a strongly nonlinear lattice with viscous dissipation ", *Physical Review E*, **75**, 021304, 2007.

[32] H Hofer and E Zehnder, "Symplectic invariants and Hamiltonian dynamics". *Birkhäuser Advanced Texts*, Birkhäuser Verlag, Basel, 1994.

[33] J Hong, "Universal power-law decay of the impulse energy in granular protectors", *Physical Review Letters*, **94**, 108001, 2005.

[34] TJR Hughes and WK Liu, "Implicit-explicit finite elements in transient analysis: Stability theory", *Journal of Applied Mechanics*, **78**, 371-374, 1978.

[35] TJR Hughes, "The Finite Element Method: Linear Static and Dynamic Finite Element Analysis". Prentice-Hall, Englewood Cliffs, NJ, 1997.

[36] DD Humphreys, RA Friesner, and BJ Berne, "A multiple-time step molecular dynamics algorithm for macromolecules", *Journal of Chemical Physics*, **98**:27, 6885–6892, 1994.

[37] KH Hunt and FRE Crossley, "Coefficient of restitution interpreted as damping in vibroimpact", *ASME Journal of Applied Mechanics*, **42**, 440–445, 1975.

[38] JA Izaguirre, S Reich, and RD Skeel, "Longer time steps for molecular dynamics", *Journal of Chemical Physics*, **110**:20, 9853–9864, 1999.

[39] P Julián, A Desages, and O Agamennoni, "High-level canonical piecewise-linear representation using a simplicial partition", *IEEE Transactions on Circuits and Systems I*, **46**:4, 463–480, 1999.

[40] C Kane, JE Marsden, and M Ortiz, "Symplectic energy-momentum preserving variational integrators", *Journal of Mathematical Physics*, **40**:7, 3353–3371, 1999.

[41] C Kane, EA Repetto, M Ortiz, and JE Marsden, "Finite element analysis of nonsmooth contact", *Comput. Methods Appl. Mech. Engrg.*, **180**:1-2, 1–26, 1999.

[42] C Kane, JE Marsden, M Ortiz, and M West, "Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems", *Int. J. Numer. Meth. Eng.*, **49**:10, 1295–1325, 2000.

[43] S Kang and LO Chua, "A global representation of multidimensional piecewise-linear functions with linear partitions", *IEEE Transactions on Circuits and Systems*, **25**:11, 938–940, 1978.

[44] G Karypis and V Kumar, "Unstructured graph partitioning and sparse matrix ordering system", Technical Report, Dept. of Computer Science and Engineer, University of Minnesota, 2007.

[45] J Katzenelson, "An algorithm for solving nonlinear networks", *Bell Syst. Tech. J.*, **44**, 1605–1620, 1965.

[46] D Khatri, C Daraio, and P Rizzo, "Coupling of highly nonlinear waves with linear elastic media", *SSN07 SPIE Sensors and Smart Structures/NDE, 16th Annual International Symposium*, San Diego, CA, 2008.

[47] HW Kuhn, "Some combinatorial lemmas in topology", *IBM J. of Research and Development*, **4**, 518–524, 1960.

[48] HM Lankarani and PE Nikravesh, "A contact force model with hysteresis damping for impact analysis of multi-body systems", *ASME J. Mech. Des.*, **112**, 369–376, 1990.

[49] AN Lazaridi and VF Nesterenko, "Observation of a new type of solitary wave in a one-dimensional granular medium", *Prikl. Mekh. Tekh. Fiz.*, **3**, 115, 1985. [English translation: *J. Appl. Mech. Tech. Phys.*, **3**, 405, 1985].

[50] RA LaBudde and D Greenspan, "Discrete mechanics - A general treatment", *Journal of Computational Physics*, **15**, 134–167, 1974.

[51] A Lew, JE Marsden, M Ortiz, and M West, "Asynchronous variational integrators", *Arch. Rational Mech. Anal.*, **167**, 85–146, 2003.

[52] A Lew, JE Marsden, M Ortiz, and M West, "An overview of variational integrators", *Finite element methods: 1970's and beyond*, CIMNE, Barcelona, 2003.

[53] A Lew, JE Marsden, M Ortiz, and M West, "Variational time integrators", *Int. J. Numer. Meth. Eng.*, **60**:1, 151–212, 2004.

[54] RG Littlejohn and M Reinsch, "Internal or shape coordinates in the n-body problem", *Physical Review A*, **52**:3, 2035–2051, 1995.

[55] RG Littlejohn and M Reinsch, "Gauge fields in the separation of ratations and internal motions in the n-body problem", *Reviews of Modern Physics*, **69**:1, 213–275, 1997.

[56] DW Marhefka and DE Orin, "A compliant contact model with nonlinear damping for simulation of robotic systems", *IEEE Trans. Syst. Man Cybern., Part A. Syst. Humans*, **29**, 566–572, 1999.

[57] JE Marsden and TS Ratiu, "Introduction to mechanics and symmetry", 2th Edition. Springer-Verlag, New York, 1999.

[58] JE Marsden and M West, "Discrete mechanics and variational integrators", *Acta Numerica* **10**, 357–514, 2001.

[59] JE Marsden, G Misiolek, J-P Ortega, M Perlmutter, and TS Ratiu, "Hamiltonian reduction by stages". Springer-Verlag, New York, 2007.

[60] KR Meyer and GR Hall, "Introduction to Hamiltonian dynamical systems and the n-body problem". *Applied Mathematical Sciences* **90**, Springer-Verlag, New York, 1992.

[61] R McLachlan and C Scovel, "A survey of open problems in symplectic integration", *Fields Inst. Commun.* **10**, 151-180, 1996.

[62] S Müller and M Ortiz, "On the Γ-convergence of discrete dynamics and variational integrators", *Journal of Nonlinear Science* **14**, 279–296, 2004.

[63] M Neal and T Belytschko, "Explicit-explicit subcycling with non-integer time step ratios for structural dynamics systems", *Computers and Structures*, **31**:6, 871-880, 1989.

[64] VF Nesterenko, "Propagation of nonlinear compression pulses in granular media", *Prikl. Mekh. Tekh. Fiz.*, **5**, 136, 1983. [English translation: *J. Appl. Mech. Tech. Phys.* **5**, 733, 1984].

[65] VF Nesterenko, "Dynamics of heterogeneous materials". Springer-Verlag, New York, 2001.

[66] JN Lin, H Xu, and R Unbehauen, "A generalization of canonical piecewise-linear functions", *IEEE Transactions on Circuits and Systems I*, **41**:4, 345–347, 1994.

[67] M Ortiz, and B Nour-omid, "Unconditionally stable concurrent procedures for transient finite-element analysis", *Comput. Method Appl. M.*, **58**:2, 151–174, 1986.

[68] A Pandolfi, C Kane, JE Marsden, and M Ortiz, "Time-discretized variational formulation of non-smooth frictional contact", *Int. J. Numer. Meth. Eng.*, **53**:8, 1801–1829, 2002.

[69] J Pesonen, "Vibratio-rotation kinetic energy operators: A geometric algebra approach", *Journal of Chemical Physics*, **114**:24, 10598–10607, 2001.

[70] MA Porter, C Daraio, I Szelengowicz, EB Herbold, and PG Kevrekidis, "Highly nonlinear solitary waves in heterogeneous periodic granular media", *Physica D*, **238**, 666–676, 2009.

[71] S Reich, "Backward error analysis for numerical integrators", *SIAM J. Numer. Anal.*, **36**, 1549–1570, 1999.

[72] A Rosas, AH Romero, VF Nesterenko, and K Lindenberg, "Observation of two-wave structure in strongly nonlinear dissipative granular chains", *Physical Review Letters*, **98**, 164301, 2007.

[73] S Sen, J Hong, J Bang, E Avalos, and R Doney, "Solitary waves in the granular chain", *Physics Reports*, **462**:2, 21–66, 2008.

[74] A Spadoni and C Daraio, "Generation and control of sound bullets with a nonlinear acoustic lens", *PNAS*, DOI:10.1073/pnas.1001514107, 2010.

[75] Y Tanaka, Y Yamazaki, and K Okumura, "Bouncing gel balls: Impact of soft gels onto rigid surface", *Europhysics Letters*, **63**:1, 146–152, 2003.

[76] Y Tanaka, "Impact of gel balls beyond the Hertzian regime", *European Physical Journal E Soft Matter*, **18**:1, 95–103, 2005.

[77] Y Tanaka, "Double-peaked impact force of very soft gel balls", *Physical Review E*, **73**, 031403, 2006.

[78] L Vergara, "Model for dissipative highly nonlinear waves in dry granular systems", *Physical Review Letter*, **104**, 118001, 2010.

[79] H Whitney, "Geometric integration theory". *Princeton University Press*, Princeton N.J., 1957.

[80] P Wriggers, "Computational Contact Mechanics". Wiley, New York, 2002.

[81] T Yanao, WS Koon, and JE Marsden, "Gyration-radius dynamics in structural transitions of atomic clusters", *Journal of Chemical Physics*, **126**, 124102, 2007.

[82] Y Zhang and I Sharf, "Validation of nonlinear viscoelastic contact force models for low speed impact ", *Journal of Applied Mechanics*, **76**:5, 051002, 2009.

[83] OC Zienkiewicz and RL Taylor, "The Finite Element Method", 5th Edition. Butterworth-Heinemann, Stonehamm, MA, 2000.